

STIC-Biotech/ChemLib

151705

mg

From: Jiang, Dong  
Sent: Tuesday, April 26, 2005 11:48 AM  
To: STIC-Biotech/ChemLib  
Subject: 10/075,846

Please search 1) SEQ ID NO:3  
2) nucleotides 574-621 of SEQ ID NO:3

-pending  
-issued & Pub.  
-commercial

Please send results on paper to Dong Jiang in REM 4D70 (mail stop REM 4C70).  
Thank you very much.

*Dong Jiang*

AU1646  
REM - 4D70  
571-272-0872  
Mail stop REM - 4C70

\*\*\*\*\*  
STAFF USE ONLY

Searcher: \_\_\_\_\_  
Searcher Phone: 2- \_\_\_\_\_  
Date Searcher Picked up: 4/27/05  
Date Completed: 4/29/05  
Searcher Prep/Rev. Time: \_\_\_\_\_  
Online Time: \_\_\_\_\_

\*\*\*\*\*  
Type of Search

NA#: 2 AA#: \_\_\_\_\_  
Interference: \_\_\_\_\_ SPDI: \_\_\_\_\_  
S/L: \_\_\_\_\_ Oligomer: \_\_\_\_\_  
Encode/Transl: \_\_\_\_\_  
Structure#: \_\_\_\_\_ Text: \_\_\_\_\_  
Inventor: \_\_\_\_\_ Litigation: \_\_\_\_\_

\*\*\*\*\*  
Vendors and cost where applicable

STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
QUESTEL/ORBIT: \_\_\_\_\_  
LEXIS/NEXIS: \_\_\_\_\_  
SEQUENCE SYSTEM: ot  
WWW/Internet: \_\_\_\_\_  
Other(Specify): \_\_\_\_\_

**This Page Blank (uspto)**

Result No.	Query			Length	DB	ID	Description
	Score	Match					
1	1521.2	92.8	1874	6	AX392950	AX392950 Sequence	
2	1503.8	91.7	1600	6	AX037565	AX037565 Sequence	
3	1002.2	61.1	1251	10	AF462147	AF462147 Mus musculus	
4	695.8	42.4	2045	5	DR5404970	DR5404970 Danio rerio	
5	659.2	40.2	1384	10	RNGRSTR	R. norvegicus	
6	656.2	40.0	1407	10	S73717	S73717 GLRA1=Inhib	
7	656	40.0	1439	9	BC074980	BC074980 Homo sapi	
8	656	40.0	1712	6	C0728075	C0728075 Sequence	
9	656	40.0	1715	9	HSLGYRA2	HSLGYRA2	
10	654.8	39.9	1359	10	RNO310837	RNO310837 Rattus norvegicus	
11	654.8	39.9	3160	10	RNIGRAA2	RNIGRAA2	
12	654.8	39.9	3865	6	AX700251	AX700251 Sequence	
13	654.8	39.9	3865	10	RNNEOGLY	RNNEOGLY	
14	654.6	39.9	1350	10	RNO310834	RNO310834 Rattus norvegicus	
15	654.6	39.9	1388	10	AY129229	AY129229 Mus musculus	
16	654.6	39.9	2125	10	RATGLYRAL	RATGLYRAL	
17	653.2	39.8	2775	10	BC056342	BC056342 Rattus norvegicus	
18	653.2	39.8	2775	10	BC068987	BC068987 Mus musculus	
19	650.2	39.6	1857	9	HSLGYRA1	HSLGYRA1	
20	650.2	39.6	1857	9	X52008	X52008 H.sapiens a	





[illegible]

Qy	1141	CAAGAAAGTCGTTTCTATATTTCCGTGGCTATGGCTTGGGCCACTGCTGCGAGGCAAGAGAT	1201
Ds	1099	CAAGAAAGTCGTTTCTATATTTCCGTGGCTATGGCTTGGGCCACTGCTGCGAGGCAAGAGAT	1158
Qy	1201	GGAGGTCGAATGGAAGGTCCTGCGATTTATAGTCCCAACTCTCAGGCCCTCTTCTAAAG	1260
Ds	1159	GGAGGTCGAATGGAAGGTCCTGCGATTTATAGTCCCAACTCTCAGGCCCTCTTCTAAAG	1218
Qy	1261	GAAGGAGAAACCAACGCGGAAACTCTACGTGGACTGAGCCAGAGAAATTGACCAATCTCC	1320
Ds	1219	GAAGGAGAAACCAACGCGGAAACTCTACGTGGACTGAGCCAGAGAAATTGACCAATCTCC	1278
Qy	1321	C-GGGCTGCTTCCCTTTCATCTTCCATCTCAATCTTCAATCTTCTACTGGGTTGTCTATAA	1379
Ds	1279	CAAGGGCGAATTCCTTTTTCATCTTCCCTCATCTTCAATCTTCTACTGGGTTGTCTATAA	1338
Qy	1380	AGTGCTATGTCAGAGATATCCACAGGCTCTGTGAATAGGGTGGGAGCTATAGAGTCC	1439
Ds	1339	AGTGCTACGGTCAGAGATATCCACAGGCTCTGTGAATAGGGTGGGAGCTATAGAGTCC	1398
Qy	1440	TGCTGTGGGCTCCTGCTTCTCCTCTGGGTTGGGCTTCTCCTCAGTTPAGACTCCATPAGG	1499
Ds	1399	TGCTGTGGGCTCCTGCTTCTCCTCTGGGTTGGGCTTCTCCTCAGTTPAGACTCCATPAGG	1457
Qy	1500	GGTTTGACAGTTCCTTCTGATCTCCACTCAGAACTTCAACTACCAGTCCCAAGCTA	1559
Ds	1458	GGTTTGACAGTTCCTTCTGATCTCCACTCAGAACTTCAACTACCAGTCCCAAGCTA	1517
Qy	1560	TGTGGGCTATATTTGCAATGCTGGCAATGGTGGGCTGTAATTTATAAGATGGCTTATCTACC	1619
Ds	1518	TGTGGGCTATATTTGCAATGCTGGCAATGGTGGGCTGTAATTTATAAGATGGCTTATCTACC	1577
Qy	1620	CTTAATAAAAAAAAAAAAAA 1640	
Ds	1578	CCTAATAAAAAAAAAAAAAA 1598	
RESULT 3			
AF462147			
LOCUS			
DEFINITION	AF462147	1251 bp mRNA linear ROD 11-FEB-2002	
		Mus musculus glycine receptor alpha 4 subunit (Glr4) mRNA,	
		complete cds.	
ACCESSION	AF462147		
VERSION	AF462147.1	GI:18448710	
KEYWORDS			
ORGANISM		Mus musculus (house mouse)	
		Mus musculus	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS		Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
TITLE		1 (bases 1 to 1251)	
		Localization of different glycine receptor isoforms in murine	
		spinal cord	
JOURNAL		Unpublished	
REFERENCE		2 (bases 1 to 1251)	
AUTHORS		Groemer, T.-W., Becker, C.-M. and Becker, K.	
TITLE		Direct Substitution	
JOURNAL		Submitted (21-DEC-2001) Biochemistry, University of Erlangen,	
		Fahrstrasse 17, Erlangen 91054, Germany	
FEATURES		Location/Qualifiers	
source		1..1251	
		/organism="Mus musculus"	
		/mol_type="mRNA"	
		/strain="B6C3/Fe"	
		/db_xref="taxon:10090"	
		/chromosome="X"	
		/tissue_type="spinal cord"	
gene		1..1251	
		/gene="Glr4"	
		/note="synonym: mmGlr4"	
CDS		1..1251	
		/gene="Glr4"	

Db	739	GTAGGCTTGGCATCACCACTCTCCACCATGACAATCAGAGCTCTGGATCCAGGCC	7398
Qy	964	TCCTTGCCTAAGTGTCCTACGTTGAAGGCAATCGACATCTGGATGGCTGTGTCTGCTC	1023
Db	799	TCCTTGCCTAAGGTGCTACGTTAAAGGCAATTGACATCTGGATGGCTGTATGCCTGCTC	858
Qy	1024	TTTGTGTCGCTGCCCTTGGTAGTAGTGTGCCATAAATTTGTTTCTCGTCAGCATAA	1083
Db	859	TTTGTATTGCTCGCTTGGTTGAGTAGTGTGCAGTCAAATTTTGTCTCGTCAGCATAA	918
Qy	1084	GAATTTCATACGACTTCGAAGAAGCGAGAGCGCCAAACGCTTGGAGGAAGATATCATCCAA	1143
Db	919	GAATTTATGAGACTTCGTAGAAGCGAGAGACGTCAGCGTATGGAAGAAGACATCAAGA	978
Qy	1144	GAAGTTCGTTTCTATTTCCGTGCTATGCTTGGGCCACTGCCCTGCAGGCAAGAGATGA	1203
Db	979	GAGAGCGCTTCTATTTTGTGTGCTATGGCTTAGGACACTGCTTACAAGCAAGGGATGA	1038
Qy	1204	GGTCCCAATGGAAGGTTCTGGCAATTTAGTCCCAACCTCCAGCCCTCTTCTAAAGGAA	1263
Db	1039	GGTCCCAATGGAAGGTTCCAGCAATTTAGTCCCAACCTCCAAACCTCTTCTAAAGAA	1098
Qy	1264	GGAGAAACACGCGGAAACTCTACGTGACTGAGCCAAAGAAATTGACACCATCTCCCGG	1323
Db	1099	GGAGAAACCATGAGGAACTCTATGTGGACCGAGCCAAAGAAATTGACACCATCTCCAGG	1158
Qy	1324	GCTGTCTTCCCTTCACTTTCCTCATCTTCAATATCTTCTAC'TGGGTGTCTATAAAGTG	1383
Db	1159	GCTGTCTTCCCTTCACTTTCCTCATCTTCAATATCTTCTATTTGGGTGTCTATAAAGTA	1218
Qy	1384	CTATGTCAGAAGATATCCACCGGCTCTGTGA	1416
Db	1219	CTAAGGTCAGAAGATATCCACGAGCACTGTGA	1251

Best Local Similarity 87.2%; Pred. No. 2.2e-264;  
Matches 1128; Conservative 0; Mismatches 123; Indels 42; Gaps 1;

RESULT 4  
DRE404970

LOCUS  
DRE404370

### DEFINITION

**ACCESSION**

**VERSION**  
**KEYWORDS**

**KEYWORDS**  
**COITRCE**

**SOURCE**  
**ORGANIZATION**

## DATA

## REFERENCE

## AUTHORS

**TTTTT**

### ATTITI

JOURNAL

## REFERENCE

## AUTHORS

**TITLE**

**JOURNAL**

**FEATURING**

## FEATURES

inos

gene

5/11/2017

10-5

CDS

23



```

/clone="GR1"
/tissue_type="spinal cord"
<1..1286
/codon_start=3
/product="glycine receptor strychnine binding subunit"
/protein_id="CAA68378.1"
/db_xref="GI:755778"
/db_xref="GOA:P07727"
/db_xref="UniProt/Swiss-Prot:P07727"
/tranlation="SKEVDAARSAPKSPSPDLKLMGRSTGYDARIPNFKGPPVN
VSCNFINFSGIAETMDYRNYFLRQNDPRLAYNEYPPDSDLDSLWKP
DUFANEKGAHFEIDYRNGVLYSIRITLLACPMDLKPNFMDVQICIM
QLESFGYNDLIFEQEGAGVADGLTPQILKEKDLRYCTKHNTGKPTCEIA
RHLERQMYVYIIPISLLIWSWIMDAAPARVGLITVLTMTQSSGS
RFLPKVSVYKIDIMWALFVFSALLEVAANFVSRQKELLERFRKRHHKDE
GEGRENFSAWGMGPACAKDGISVKGANNTNTPAPSPKPEENRKLFIQRAKK
IKLSRIGFPMAPLIPNMFYIYKIVREDVHNK"
mat_peptide
21..1283
misc_feature
561
/product="glycine receptor strychnine binding subunit"
/notes="c/t transition"
ORIGIN
Query Match 40.2%; Score 659.2; DB 10; Length 1384;
Best Local Similarity 72.1%; Pred. No. 6.3e-170;
Matches 954; Conservative 0; Mismatches 313; Indels 57; Gaps 5;
QY 89 AGGAAGTCAATTCGAAACCCAGGGTCCAGCCCATGTCCTCTGATTTCCCTAGACA 148
DB 7 AGAGGTGACGCTGCGCGCTCGACCCCAAGGCTATGACCCCTCGGACTTCTCGGATA 66
QY 149 AACTTATGGGGGAAACATCTGGATATGATGCCAGGATTCGGGCCAATTTTAAAGGCCAC 208
DB 67 AGCTTATGGGAGGACTTCTGGGTATGATGCCAGGATCAGACCACTTTAAAGGTCCTC 126
QY 209 CGGTGAACGTGACCTCAACATCTTCATCAACAGTTTCCAGTCCGTCAACAAGACACAA 268
DB 127 CTGTGAACGTGAGTGCACATCTTCATCAACAGCTTTGGTCTATATCGCGGAGACAA 186
QY 269 TGGACTACCGGTGATGCTCTTGGCGMACAGTGGATGACCCACCGCTGCTCTACC 328
DB 187 TGGACTACAGGGTCAACATCTTCTGAGGAGCAGTGGAGAACGCCCTGCTGCGCTACA 246
QY 329 GAGATATCTGATGACTCTCTGGACCTCGATCCCTCCATGCTGGACTCTATCTGGAAGC 388
DB 247 ATGAATATCTGACGACTCTGACCTTGACCCATCCATGTTGGATTCCATCTGGAAGC 306
QY 389 CAGACCTCTTCTTGTGATATGAGAAAGGGCCCAACTTCCATGAGGTGACACCGACAA 448
DB 307 CTGACTTGTTCTTGGCAATGAGAGGGGGCCCACTTCCACGAGATCACCACGACAA 366
QY 449 AGTTACTCGGCATCTTCAAGATGGGAATGCTGTACAGCATCAGGCTGACCCCTCATTT 508
DB 367 AGTGTCTGAGAATCTCCGGAAGCGGACGCTCTCTACAGCATCAGAACACCCCTGACTC 426
QY 509 TGTCTGCTGATGAGACCTCAAGAACTTCCCATGGACATCCAGACGTCGACGATCAGC 568
DB 427 TGGCTTCCCCCATGACCTCAAGAAATTTCCGATGACGCTACAGACATGATATCATCA 486
QY 569 TTGAGAGCTCATCCATFACTCTGAGCGCTCTGCGCTCTGTCTGCTGCTGCTGCTGCT 628
DB 487 TGGAAAGCT-----ITGGTTATA 504
QY 629 CCATGAAGACCTCGTGTGAGTGGCTGGAAGATGCTCTCTGCTGCTCAAGTGGCTGAG 688
DB 505 CCATGAAGACCTCATCTTTGAGTGGC---AAGCAAGAGCTGTCGAGGTGGCAGATG 561
QY 689 GGCTGACTCTGCCCCAGTTTATCTTGGGATGAGAGGATCTAGGCTGTGTGACCAAGC 748
DB 562 GACTGACCTGCTCAGTTTATCTTGAAGAGAGAGGATCTGAGATCTGACACCAAGC 621
QY 749 ACTAACAACAGGAAATTCACCTGATCGAGTAAAGTTTACCTGGAAACGGCAGATGG 808
|||||

```

```

DB 622 ACTACAACACAGGTAAATTCACCTGCATTGAGGCCGGAATCCACCTGGAAACGGCAGATGG 681
QY 809 GCTACTATCTGATTCAGATGTATACATCCAGCCTTACTCATCTGCTATCTCTGCTGCTGCT 868
DB 682 GCTACTATCTGATTCAGATGTATACATCCAGCCTTACTCATCTGCTATCTCTGCTGCTGCT 741
QY 869 CTTCTGATCAACATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 928
DB 742 CTTCTGATCAACATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 801
QY 929 TCACATGACCAACAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 988
DB 802 TCACATGACCAACAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 861
QY 989 AGGCAATGACATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1048
DB 862 AGCTATGACATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 921
QY 1049 ATGCTGCCATAAATTTGTTCTCTGTCAGCATPAAAGAAATTCATACGACTTTCGAAAGAGGC 1108
DB 922 ATGCGCTGTCAACTTTGTTGCTCCGCAACACAGGAATC-----CTTGATTTAGGA 975
QY 1109 AGAGGGCCCAACGCTTGGAGGAGATATCATCCAGAAAGTCTGTTCTATTTTCCGTGGCT 1168
DB 976 GGAAGCGGCGACATCACAGGATGATGAGGGTGGAGAGGCCGCTTCAACTTCTCCGCT 1035
QY 1169 ATGGCTTGGGCC---ACTGCTGAGGAGAGAGATGAGGTCCAAATGGAAGTCTTGGA 1225
DB 1036 ATGGATGGGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1095
QY 1226 TTTATAGTCCCAACCTCCAGCCCT---CTTCTAAGGGAAGAGAAACCCAGCGGAAC 1282
DB 1096 ACAACAACACACGACCCCGCTCTGCAAGTCCCAAGTCCCGGAGGAGATGCGGAAC 1155
QY 1283 TCTAGTGGATGAGGAGGAGAGAGATGACACATCTCCGGGCTGCTTCCCTTTCACCTT 1342
DB 1156 TCTTATCAGAGAGCCAGAAAGATCGAATAATATCTGCAATCGGTTTCCCGCTGCT 1215
QY 1343 TCTCATCTTCAATCTTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1402
DB 1216 TCTCATCTTCAATCTTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1275
QY 1403 ACCA 1406
DB 1276 ACA 1279

RESULT 6
S73717
LOCUS
DEFINITION
S73717
VERSION
S73717.1
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REMARK
FEATURES
source
1..1407
/organism="Mus sp."
/mol_type="mRNA"

```

S73717 1407 bp mRNA linear ROD 17-JUL-2000  
 GLRA1=Inhibitory glycine receptor alpha 1 subunit [mice, A/HeJ,  
 brainstem, spinal cord, mRNA, 1407 nt].  
 S73717  
 S73717.1 GI:765206  
 Mus sp.  
 Mus sp.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 1407)  
 Ryan, S.G., Buckwalter, M.S., Lynch, J.W., Handford, C.A., Segura, L.,  
 Shiang, R., Wasmuth, J.J., Camper, S.A., Schofield, P. and O'Connell, P.  
 A missense mutation in the gene encoding the alpha 1 subunit of the  
 inhibitory glycine receptor in the spasmodic mouse  
 Nat. Genet. 7 (2), 131-135 (1994)  
 95004575  
 7920629  
 GenBank staff at the National Library of Medicine created this  
 entry [NCBI gi|56492| from the original journal article.  
 Location/Qualifiers  
 1..1407  
 /organism="Mus sp."  
 /mol\_type="mRNA"

```

/db xref="taxon:10095"
/tissue type="brainstem and spinal cord"
/notes="A/HeJ"
1. .1407
/genes="Giral"
58. .1407
/genes="Giral"
/notes="alternatively spliced; conceptual translation
differs from that in published reference; Giral protein;
GlyR alpha 1"
/codon start=1
/product="inhibitory glycine receptor alpha 1 subunit
short form"
/protein_id="AAB32157.2"
/db_xref="GI:9247215"
/translation="MYSFNTLRYLWETIVFFSLAASKAEAAARSPKMSDFLDK
LMRGTYDARIIRPNFPGPVNSCNIFINSFGSIAETMDYRNVIFLQWNPDLRLA
YNEYPDSDLDPSMDLSIWKPDLPFANEKGAFHEITDNLKLRIRSNQNVLSIRI
TLTACPMDLKPNFMDVQTCIMOLESGFTMNDLIFEWEOGAVQVADGLTLPQFLLK
EEKDLRYCTKHNTGKFTCIEARPHLEROMGYLIOMYIPSLIVILSWISWINDA
APARVLGITTIVLTMTQSSGRASLPKYSVNAIDIMRAVCLLFFLSALLEYAAVNF
VSHQKELLFRFRKRHHDDGEGGRFNPAYMGPGPACLQAKDGLSVKGNNTTN
PPAPSKSPFEMERKLIQRAKKIDKISIRIGFPMAFLIFNMFYIYKIVREDVHNK"

ORIGIN
Query Match 40.0%; Score 656.2; DB 10; Length 1407;
Best Local Similarity 72.6%; Pred. No. 4.3e-169;
Matches 942; Conservative 0; Mismatches 298; Indels 57; Gaps 5;

QY 116 CCAGGCCATGTCCTCCCTCTGATTTCTAGACAAATTTATGGCGGCAACATCTGGATATG 175
DB 155 CCAAGCTATGTACCTCTGGACTTCTCTGATTAAGTCATCTGGGAGGACTTCTGGGTATG 214
QY 176 ATGCCAGGATTCGGCCCAATTTTAAAGCCCAACCGGTGAACGTGACCTGCAACATCTTCA 235
DB 215 ACGCCAGGATCAGGCCCACTTTTAAAGTCTCTCTGTGAACGTAAAGTTGCAACATCTTCA 274
QY 236 TCAACAGTTTCAGCTCGGTCCAGAACCAATGAGTACCTACCGGTGAATGCTCTCTGTC 295
DB 275 TCAACAGTTTCAGTTCATCGCTGAGACCAACATGAGTATAGGCTCAACATCTCTCTGA 334
QY 296 GGCAACAGTGAATGATCCACGCTGTCTTACCGAGAAATATCTGTGATGACTCTCTGACC 355
DB 335 GGCAAGAGTGAACGACCCCGCTGTGGCTTCAATGATACCTGATGACTCTCTGACC 394
QY 356 TCGATCCCTCCATGCTGAGCTATCTTGAAGCCAGACCTCTCTTTTGTGAATGAGAAAG 415
DB 395 TTGACCCATCTATGTTGATTCATCTGGAAGCTGACTTGTCTTTTGCCAAATGAGAAG 454
QY 416 GGCCCACTTCCATGAGTGACCAACGGAACAAAGTTACTGGCGATCTTCAAGAAATGGGA 475
DB 455 GGCCCACTTCCAGAAATCACCACGGAACAAACTGTCTGAAGAAATCTCCCGAAATGGCA 514
QY 476 ATGTGCTGTACAGATCAGCTGACCTCTATTTTGTCTGCTGATGAGCTCAAGAAT 535
DB 515 ATGTCTCTACAGATCAGAAATCACCTCTGAGCTGGCTGCCCATGAGCTCAAGAAAT 574
QY 536 TCCCATGGAATCCAGACGTGACAGTGCAGCTTGAAGAGCTCATCCATCTGCGAGCC 595
DB 575 TCCGATGATGTACAGACGTGTATCATGCACTCAACTCGAAAGCT----- 616
QY 596 CTCTGCCATCTGTCTACTTTCAGTTGGCTTACACATGAAAGACCTGCTGCCCCAGTTTCTTGC 715
DB 617 -----TTGGATATACCATGAACGACCTCATCTTTGAGTGGC 652
QY 656 TGAAGATGCTCTGCTGTCCAAAGTGGCTCAGAGGCTGACTCTGCCCCAGTTTCTTCTG 715
DB 653 AGGAGCAAG---GAGCTGTGACAGTGGCGACAGGACTGACCCCTGCTCTGCTTTTCTGA 709
QY 716 GGGATGAGAGGATCTAGGCTGTGTGTAACCAAGCACTACAAACAGAGGAAATTCACCTGCA 775
DB 710 AGGAGAGAGGACCTGAGATAGTGTGACCAAGCACTACAAACAGAGTAAATTCACCTGCA 769

```

```

QY 776 TCGAGCTAAAGTTTTCACCTGGACGCGAGATGGCTACTATCTGATTTCAGATCTACATCC 835
DB 770 TCGAGGCCCGATTTCCACCTGGAGCGGAGATGGGCTACTACCTGATCCAGATGTACATCC 829
QY 836 CCAGCCTACTCATCTGTCATCTCTCTCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 895
DB 830 CCAGCCTGCTCATCTGTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 889
QY 896 CTGCCCTGTGGGCTGGGATCACCACCGTGTCTACCATGACCAACCCAGAGCTCTGGCT 955
DB 890 CAGCTGTGTGGGCTGGGATCACCACAGTGTCTACCATGACCAACACAGAGCTCTGGCT 949
QY 956 CCGGGCTCTTTTGGCTTAAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1015
DB 950 CCGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1009
QY 1016 GTCTGCTCTTTTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1075
DB 1010 GCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1069
QY 1076 AGCATAAAGATTCATACGACTTGAAGAGGAGGAGGCGCCCAACGCTTGGAGAGAGATA 1135
DB 1070 AACACAAGGA-----ACTGCTTCGATTTAGGAGGAGAGGCGGACATCAAGGATGATG 1123
QY 1136 TCATCCAAAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1192
DB 1124 AGGCTGGAGAGGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1183
QY 1193 CAAAGATGGAGTCCCAATGGAAGTCTCTGGAATTTATAGTCTCTCTCTCTCTCTCTCT 1249
DB 1184 CCAAGATGGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1243
QY 1250 CTCTTCTAAGGAGAGGAGAAACCAACCGGAACTCTACGTGGACTGAGCCAGAGAAATG 1309
DB 1244 CGCCATCAAGTCCCGGAGGAGATCGGGAACCTCTTCTATCCAGAGAGCCAGAGATCG 1303
QY 1310 ACACCATCTCTCGGGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1369
DB 1304 ACAAGATATCTCGCATCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1363
QY 1370 TTGTCTATAAAGTCTATGTCAGAGAGATATCCACCA 1406
DB 1364 TCATCTATAAGATCTCTCGGAGAGAGATGTCCACAA 1400

BC074980 1439 bp mRNA linear PRI 04-AUG-2004
Homo sapiens glycine receptor, alpha 1 (startle
disease/hyperkplexia, stiff man syndrome), mRNA (cDNA clone
MGC:103909 IMAGE:30915298), complete cds.
BC074980
VERSION BC074980.2 GI:50959672
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1439)
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L.,
Scheetz T.E., Brownstein M.J., Ustin T.B., Toshiyuki S.,
Carninci P., Frange C., Raha S.S., Loquellano N.A., Peters G.J.,
Abramson R.D., Mullaly S.J., Bosak S.A., McSwan P.J.,
McKernan K.J., Hale S.J., Gunaratne P.H., Richards S.,
Worley K.C., Hale S.J., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Y.,
Sanchez A., Whitting M., Madan A., Young A.C., Shevchenko Y.,

```

Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,  
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
 Butterfield, Y.S., Krzyzanski, M.I., Skalska, U., Smalhus, D.E.,  
 Scherch, A., Schein, J.E., Jones, S.J., and Marra, M.A.  
 Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
 12477932  
 2 (bases 1 to 1439)  
 Director MGC Project.  
 Direct Submission  
 Submitted (25-JUN-2004) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 On Aug 4, 2004 this sequence version replaced gi:49901657.  
 Contact: MGC help desk  
 Email: [cgapsb-r@mail.nih.gov](mailto:cgapsb-r@mail.nih.gov)  
 Tissue Procurement: Genome Sequence Centre, British Columbia Cancer  
 Center  
 cDNA Library Preparation: British Columbia Cancer Research Center  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Genome Sequence Centre,  
 BC Cancer Agency, Vancouver, BC, Canada  
 info@bcgsc.bc.ca  
 Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,  
 Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth  
 Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,  
 Kim Macdonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,  
 Teika Olson, Diana Palmquist, Anca Petrescu, Anna Lilia Prabh,  
 Parvaneh Saeedi, JR Santos, Angeliue Schnerch, Ursula Skalska,  
 Duane Smalhus, Jeff Stott, Miranda Tsai, George Yang, Jacquie  
 Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
 Series: IRBU Plate: 2 Row: C Column: 11.  
 Location/Qualifiers  
 1..1439  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="MGC:103909 IMAGE:30915298"  
 /tissue\_type="Pooled tissue, PCR rescued clones"  
 /clone\_lib="NIH MGC\_245"  
 /notes="Vector: PCR4 Topo TA"  
 1..1439  
 /genes="GLR1"  
 /notes="synonym: STHE"  
 /db\_xref="LocusID:2741"  
 /db\_xref="MIM:138491"  
 42..1391  
 /genes="GLR1"  
 /codon\_start=1  
 /product="glycine receptor, alpha 1"  
 /protein\_id="AAH74980.1"  
 /db\_xref="GI:49901658"  
 /db\_xref="LocusID:2741"  
 /db\_xref="MIM:138491"  
 /translation="MYSPNTRLRLYLWETIVFFSLAASKEAARSAPKPMSPDFLDK  
 LMGRTSGDYARPNPKGPPVNVSCNIFNSGFIABTTMDYRVNIFLRQWNPRLA  
 VNEYPDSDLDPSMDSITWPKPLPFANEKGAFHEITTDNKLRLSRNGNVLYSIRI  
 TLTLACPMKLPNDVQTCIMQLESFGYTMNDLIFEMQEGAVQVADGLTLQPFLK  
 EKDLRYCTKHTNGFTICIEARFLEROMGYLIQMYIPSLVILSWISFWINDA  
 APARVLGTLTVLTITQSSGRSLPYSYKVIDIWMVCLLFVFSALLEYAANF  
 VSRQKELLRRFRKRHREDEAGEFNFSAYMGPCLOAKQDGLSVKGNNSNTN  
 PPPAPSKSPENRKLFIQRAKIDKISRGFPMFLFNFWIYIIVRREDVHNQ"

Query Match 40.0%; Score 656; DB 9; Length 1439;  
 Best Local Similarity 71.9%; Pred. No. 4.8e-169;  
 Matches 952; Conservative 0; Mismatches 315; Indels 57; Gaps 5;

QY 89 AGGAAGTCAAATCTCGAACCACCAAGGGGTCCAGCCCATGTCCCCCTCTGATTTCTTAGACA 148  
 |||||  
 DB 112 AGGAGCTGAAGCTGCTCGCTCGGACCAACGCTTATGTACCCCTCGATTTCTTGGATA 171  
 |||||  
 QY 149 AACTTATGAGGGGGAACATCTGATATGATGACAGATTCGGCCCAATTTTAAAGGCCAC 208  
 |||||  
 DB 172 AGCTAATGGGAGAACCTCCGGATATGATGCCAGGATCAGGCCCAATTTTAAAGGTCCCC 231  
 |||||  
 QY 209 CGGTGAAGTGAACCTGACACATCTTCATCAACAGTTTTCAGCTCCGTCACCAAGACCA 268  
 |||||  
 DB 232 CAGTGAAGCTGAGCTGCAACATTTTCAACAGCTTTGTTTCCATTTGCTGAGACCAACA 291  
 |||||  
 QY 269 TGGACTACCGGTGAATGTCTTCTGCGGCAACAGTGAATGACCCACCCCTGTCTCTACC 328  
 |||||  
 DB 292 TGGACTATAGGGTCAACATCTTCTCGGCGCAGCATGGACGACCCCGCTGGCTATA 351  
 |||||  
 QY 329 GAGATATCTGATGACTCTCTGGAACCTGATCCCTCAATGCTGAGTCTATCTGGAAGC 388  
 |||||  
 DB 352 ATGAATACCTGACGACTCTCTGGACCTGGACCCATCCATGCTGGACTCCATCTGGAAC 411  
 |||||  
 QY 389 CAGACTCTCTTCTTCTATGAGAAAGGGGCAACTTCCATGAGGTGACCGGACCAACA 448  
 |||||  
 DB 412 CTGACTGTCTTCTTGGCAACGAGAAAGGGGGCCACTTCCATGAGATCACCAAGACAA 471  
 |||||  
 QY 449 AGTTACTGCGCATCTTCAAGAAATGGGAATGTCTGTATACAGCATCAGGCTGACCCCTCATTT 508  
 |||||  
 DB 472 AATTGTAAGGATCTCCCGGATGGGAATGTCTCTACAGCATCAGATCACCCCTGACAC 531  
 |||||  
 QY 509 TGTCTGCTGATGAGCACTCAAGAACTTCCCATGAGACATCCAGAGTGCAGATGCAGC 568  
 |||||  
 DB 532 TGGCTGCCCCATGAGCTTGAAGAAATTTCCCATGATGTCAGATCATATCATGCAAC 591  
 |||||  
 QY 569 TTGAGAGCTCATCCATCTCTGCGACCCCTCTGCCATCTCTGTCTCAGTTGGGTACA 628  
 |||||  
 DB 592 TGGAAAGCT-----TTGGATATA 609  
 |||||  
 QY 629 CCATGAAGACCTCGTGTGAGTGGTGGAAAGATGCTCCTGTGTCTCAAGTGGTGCAGG 688  
 |||||  
 DB 610 CGATGATGACCTCATCTTTGAGTGGCAGAACAGS---GAGCCGTGCAGGTAGCAGATG 666  
 |||||  
 QY 689 GGCTGACTGCTGCCCGCACTTTATCTTGGGGATGAGAGGATCTAGGCTGTTGTACCAAGC 748  
 |||||  
 DB 667 GACTAACTCTGCCCGCACTTTATCTTGAAGAAAGAGAGGACTTGAGATATCTGCACCAAGC 726  
 |||||  
 QY 749 ACTACAACACAGGAAATTCACCTGCATCGAGTAAAGTTTTCACCTGGAAACGCGAGATG 808  
 |||||  
 DB 727 ACTACAACACAGTAAATTCACCTGATGAGGCCCGGTTCCACCTGGAGCGCAGATGG 786  
 |||||  
 QY 809 GCTACTATCTGATTCAGATGTATACATCCCGCAGCTTACTCATCTGCTCATCTCTGCGGTCT 868  
 |||||  
 DB 787 GTTACTTACTGATTCAGATGTATATCCCGACCTTGTCTCATTTGCTCTCATGATCT 846  
 |||||  
 QY 869 CTTCTGAGATCAACATGATGCTGCCCTGCGCGCTGTGGGCTGGGCATCACACCGTGC 928  
 |||||  
 DB 847 CTTCTGAGATCAACATGATGCTGCACTGCTCGTGTGGGCTTAGGCATCACCACTGTGC 906  
 |||||  
 QY 929 TCACCATGACACACAGAGCTCTGCTCCCGGCGCTCTTTGCTAAGGTGCTCTACGTGA 988  
 |||||  
 DB 907 TCACCATGACACACAGAGCTCTGCTCGGCTCTCGAGCATCTCTGCGCAAGGTGCTCTATGTA 966  
 |||||  
 QY 989 AGGCAATCCACATCTGAGTGGCTGTGTGCTCTCTTTGTTGCTGCTGCTGCTGCTGAGT 1048  
 |||||  
 DB 967 AAGCCATTGACATTTGGATGGCAGTTTGGCTGCTCTTTGTTGCTCAGCCCTATTAGAAT 1026  
 |||||  
 QY 1049 ATGCTGCCATAAATTTTGTCTCGTCAAGATAAAGAAATTCATACGATTCGAGAAAGGC 1108  
 |||||  
 DB 1027 ATGCTGCCCTTAACTTTGTGTCGCGCAACATAAGAGAGTGTCTCGATTCAGAGAGAGC 1086  
 |||||  
 QY 1109 AGAGCGCCCAACGCTTGGAGGAAGATATCATCAAGAAAGTCTTTCTTATTTTCGCTGGCT 1168  
 |||||  
 DB 1087 GGAGATCATCAAGAGGATGAAGCT-----GGAGAGCGCGCTTTAACTTCTCTGCT 1140  
 |||||





DEFINITION	H. sapiens alpha-1 strychnine binding subunit of inhibitory glycine receptor mRNA.
ACCESSION	X52009
VERSION	X52009.1 GI:31850
KEYWORDS	glycine receptor; inhibitory glycine receptor; strychnine binding.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	Grenningloh, G., Schmieden, V., Schofield, P.R., Seeburg, P.H., Siddique, T., Mohandas, T.K., Becker, C.M. and Betz, H.
TITLE	Alpha subunit variants of the human glycine receptor: primary structures, functional expression and chromosomal localization of the corresponding genes
JOURNAL	EMBO J. 9 (3), 771-776 (1990)
MEDLINE	90183975
PUBLISHED	2155780
FEATURES	<p>Location/Qualifiers</p> <p>1..1715</p> <p>/organism="Homo sapiens"</p> <p>/mol_type="mRNA"</p> <p>/db_xref="taxon:9606"</p> <p>/chromosome="X"</p> <p>/map="Xp21.2-p22.1"</p> <p>/tissue_type="brain"</p> <p>/dev_stage="fetal"</p> <p>1..1715</p> <p>/gene="GLYRA2"</p> <p>297..1646</p> <p>/gene="GLYRA2"</p> <p>/notes="strychnine binding alpha-1 subunit"</p> <p>/codon_start=1</p> <p>/product="inhibitory glycine receptor"</p> <p>/protein_id="CAA36258.1"</p> <p>/db_xref="GI:31851"</p> <p>/db_xref="GOA:P23415"</p> <p>/db_xref="UniProt/Swiss-Prot:P23415"</p> <p>/translation="MYSENTLRLSLGAIIVFSLAASKEAARSATKPMSPDFLDKLMGRSGYDARPNPKGPPVNVSCNIFINSFSGIAETTDYRVNIFLQWNRDLA YNEYPDSDLOLPSMDSLWKPDLPFANEKGAFHEITTDNKLRLISRNGLVYSIRI TULACPMDLNPFNDPQCTIMQLESFGYTMNDLIPFWEQGAQVADGLTPQFLTK EEKDLRYCTKHYNTGFTCTEARPHLERQMYLIQMYIPSLIIVLSWISVINMDA APARVGLGTTVLMTTQSSGRSLPKVYVKAIDIIMAVCLLFFVSLLEVAANF VSRQKELLRFERKRHHKDEAGEGRFNFSVAGMGPACLOAKDGLSVKGANNSTN PPAPSKSPEEMRKLFIQRAKKIDKISRIGFPMFLIFNWFYIIVKIVREDVHNQ"</p>
gene	
CDS	
Query Match	40.0%; Score 656; DB 9; Length 1715;
Best Local Similarity	71.9%; Pred. No. 5e-169;
Matches	952; Conservative 0; Mismatches 315; Indels 57; Gaps 5;
QY	89 AGAAGTCAATCTGNAACCAAGGGTCCAGCCCATGTCCTCTGATTTCTTACACA 148
DB	
QY	367 AGGAGGCTGAAGCTGCTCGCTCCGCAACCAAGCCTATGTACACCTCGGATTTCTGGATA 426
DB	
QY	149 AACTTATGGGCGCAACATCTGGATATGATGCCAGGATTCGGCCCAATTTTAAAGGCCAC 208
DB	
QY	427 AGCTAATGGGGAGAACCTCCGGATATGATGCCAGGATTCAGCCCAATTTTAAAGGTCCTC 486
DB	
QY	209 CCGTGAACGTGACCTGCAACATCTTCAACACAGTTTCAGCTCCGTCACCAAGACCAACA 268
DB	
QY	487 CAGTGAACGTGAGCTGCAACATTTTCATCAACAGCTTTGGTTCCATTTGCTGAGACACCA 546
DB	
QY	269 TGACATACCGGGTGAATGTTCTTGGGGCAACAGTGGGAATGACCAACGCTGCTCTTACC 328
DB	
QY	547 TGGACTATAGGGTCAACATCTTCTCGCGGCAAGTGAAGCAACCCCGCTGCGCTATA 606
DB	
QY	329 GAGAAATATCTGATGACTCTCGACCTCGATCCCTCCATGCTGGAATCTATCTGGAAGC 388
DB	
QY	607 ATGAATACCTGACGACTCTCTGGACCTGGACCCATCCATGCTGGACTCCATCTGGAAC 656
DB	
QY	389 CAGACCTCTTTTGTGCTAATGAGAAAGGGGCCAACTTCCATGAGGTGACCAACGCAACA 448
DB	

ORIGIN

RESULT 10



RNO310837	RNO310837	1359 bp	mRNA	linear	ROD 03-APR-2001
LOCUS	Rattus norvegicus mRNA for glycine receptor alpha 2 precursor.				
DEFINITION	AJ310837				
ACCESSION	AJ310837.1	GI:13548660			
VERSION	glycine receptor alpha 2 precursor.				
KEYWORDS	Rattus norvegicus (Norway rat)				
SOURCE	Rattus norvegicus				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
REFERENCE	1	Beato, M., Groot-Kormelink, P.J., Colquhoun, D. and Sivilotti, L.G.			
AUTHORS	Concentration dependence of single channel currents through rat recombinant alpha 1 glycine receptors				
TITLE	Unpublished				
JOURNAL	2 (bases 1 to 1359)				
REFERENCE	Groot-Kormelink, P.J.				
AUTHORS	Direct Submission				
TITLE	Submitted (30-MAR-2001) Groot-Kormelink P.J., Department of Pharmacology, The School of Pharmacy, 29/39, Brunswick Square, London, WC1N 1AX, UNITED KINGDOM				
JOURNAL	Location/Qualifiers				
FEATURES	1..1359				
source	/organism="Rattus norvegicus"				
	/mol_type="mRNA"				
	/strain="Sprague-Dawley males"				
	/db_xref="taxon:10116"				
	/tissue_type="whole brain"				
	/clone_lib="rat brain marathon-ready cDNA"				
	/dev_stage="8-12 weeks"				
CDS	1..1359				
	/function="ion channel"				
	/codon_start=1				
	/evidence=experimental				
	/product="glycine receptor alpha 2 precursor"				
	/protein_id="CAC35981.1"				
	/db_xref="GOA:P22771"				
	/db_xref="GI:13548661"				
	/translation="MNQLVNIILTAFLPFLGNTNHPFAFKDHSRSGKPSQTLSP SFLDLKMGRTGYDARIIPNKPVPVNTCNIFINSFGSVETMDYVNIFLRQOV NLSRLAYSYPDDLSDLDPSMLDSIWKDLPFRANEGANFHDVTDNKLRLSKGKV LYSIRUTLSLCPMDLNPFMDVQTCMQLESFGYTMNDLIFEWLSDGPVQVABGLTL POFILEKEBELGYCTKHYNTGKTFIEVFLEROMGYLIQIWPISLLIVILSWVSF WIMDAAPARVALGITVTLTMTTQSSSRASLPKVSIVKAIIDIMWAVCLLFLFAALLE YAAVNFVSRQHKFELRLRRKRONKEEDVTRRESFNFGYMGHCLQVKGDTAVKAT PANLPQPKDADAIKKFPVDRAKRIDTISRAPPFLAFLIFNIYFIWITYKIIRHEDVH KK"				
ORIGIN	Query Match	39.9%; Score 654.8; DB 10; Length 1359;			
	Best Local Similarity	70.8%; Pred. No. 1e-168;			
	Matches	938; Conservative 0; Mismatches 332; Indels 54; Gaps 3;			
QY	84	AAAAGAGGAGTCAAACTCTGGAACCAAGGGGTCCAGCCCATCTCCCTCTGATTTCTT	143		
DB	84	AGACCATGACTCCAGGTCTGGAAAACATCCCTCGCAGACCCCTGCTCTCTCAGATTTCTT	143		
QY	144	AGACCAACTTATGGGGCGACATCTGGATATGATGCGAGGATTCGGCCCAATTTTAAAGG	203		
DB	144	GGATAAACAATGGGAAGGACATCAGGATATGATGCAAGAAATCAGGCCAAATTTTAAAGG	203		
QY	204	CCACCCGCTGAACGTGACCTGCAACATCTTCATCAACAGATTTTCAGCTCCGTCACCAAGAC	263		
DB	204	TCCTCAGTAAAGCTTACTTGCAATATTTTATCAACAGATTTTGGATCGGTACAGAAC	263		
QY	264	CACAATGGACTACCGGGTGAATCTCTTTTCGGGCAACAGTGAATGACCCAGCTGTTC	323		
DB	264	CACCATGGACTACCGAGTGAACATTTTCTTGAGACAGCAGTGAACGATTCACGGCTGGC	323		
QY	324	CTACCCAGAAATATCCTGATGACTCTCTGGACCTCGATCCCTCCATCGTGGACTCTATCTG	383		
DB	324	ATACAGTAGTACCAGATGATTCCCTGGATTTGGATCCATCGATCTGGATTCGATTTG	383		
QY	384	GAAAGCAGACCTCTCTTTCTTAATGAGAAAGGGGCAAACTTCATGAGGTGACCAAGGA	443		
DB	384	GAAACAGATTTGTTCTTTGGCCAAATGAGAAAGGAGCAATTTCCATGATGTACCACTGA	443		
QY	444	CAACAGTTACTGGCATCTTCAAGAAATGGGAATGTGTGTACAGCATCAGGCTGACCTT	503		
DB	444	TAAACAAGTTGCTGGGATTTCCAAAATAGGCAAGTGCTCTACAGTATTAGACTCACCTT	503		
QY	504	CATTTGTCTCTGCTGATGGACCTCAAGAACTTCCCATGGACATCCAGACGTGCAACGAT	563		
DB	504	GACTTATCTCTGCCATGACCTGAAGAACTTCCAAATGGATGTCCAGACCTGTACAAAT	563		
QY	564	GCAGCTGAGAGTCAATCCATCTCTGAGCCCTCTGCCATCTCTGTCACTTTCACTTTG	623		
DB	564	GCAGCTGGAGATT-----TTGG	581		
QY	624	CTACACCATGAAGACCTCGTCTTTGAGTGGCTGGAAAGATGCTCTCTGCTGCCAAGTGC	683		
DB	582	GTACACCATGAATGACCTGATATTTGAGTGGTTAAGTGATGGTCT---CAGTACAAGTGC	638		
QY	684	TGAGGGCTGACTCTGCCCCAGTTTATCTTGGGGATGAGAAAGATCTAGGCTGTGTATC	743		
DB	639	TGAGGACTCACCTGCTCAGTTTATTTTGAAGAGAGAGAACTTGGCTATTGTCAC	698		
QY	744	CAAGCACTACAACACAGGAAATTCACCTGCATCGAGGTAAAGTTTCACTGGAAACGCA	803		
DB	699	AAAGCATTAACAACACTGGCAAGTTTACCTGCAATGAGGTCAAGTTTCACTGGAAACGCA	758		
QY	804	GATGGCTACTATCTGATTCAGATGTACATCCAGCCTACTCATCGTCTCTCTGCTCTG	863		
DB	759	GATGGCTATTATTTGATCCAGATGATATCCAGCCTGCTGATAGTCAATTTGTGCTG	818		
QY	864	GGTCTCTCTGGATCAACATGATGCTGCCCTGCTGCTGGGCTTGGGCTTGGGCTCAACAC	923		
DB	819	GGTCTCTCTTTGGATTAACATGATGATGATGATGATGATGATGATGATGATGATGATG	878		
QY	924	CGTGTCTACCATGACACCCAGAGCTCTGGCTCCCGGGCTCTTTGCCCTTAAGGTGCTTA	983		
DB	879	CGTCTCTGACATGACTACCCAGAGTTTCAGGTTCAGGGCTCTCTGCCAAGGTCTCAT	938		
QY	984	CGTGAAGGCAATCGACATCTGATGCTGTGTGCTGCTCTTTGTTTTCGCTGCTGCTT	1043		
DB	939	TGTGAAGCAATGACATCTGGATGSCAGTGTGCTCTGTTTGTGTTTGTGCTGCTTACT	998		
QY	1044	GGAGTATGCTGCCATAAAATTTTGTCTCTGATGATGATGATGATGATGATGATGATGATG	1103		
DB	999	GGATATGACGAGTGAATTTTGTCTCCAGGCAACAGAGGATTCCTTCTCTCGGAG	1058		
QY	1104	AAGGCAGAGCGCCCAACGCTTGGAGGAAGATATCATCCAGAAAGTCTGTTTCTTATTTCCG	1163		
DB	1059	ACGACAGAAGAGGAGCAATTAAGGAAGATGTTACTCTGTAAGGCGTTTAACTTCA	1118		
QY	1164	TGGCTATGCTTGGGCACTGCTGAGGCAAGAGATGAGGTCCAAATGGAAGTTCG	1223		
DB	1119	CGGTTATGGATGGGTGCTGCTGCTCCAGTGAAGATGATGATGATGATGATGATGATGATG	1175		
QY	1224	CATTTATAGTCCCAACCTCCAGCCCTCTTCTTAAGGGAAGGAGAAACCAACGGGAACT	1283		
DB	1176	-----ACCTGCAACCCACTTCCGCAACCCCAAAAGATGATGATGATGATGATGATGATG	1229		
QY	1284	CTACGTGACTGAGCCCAAGAGAAATGACACCACTCTCCCGGGCTGTCTTCCCTTTCACTTT	1343		
DB	1230	GTTTGTGGATCGGGCAAAAAGAAATCGACACCATATCTCGAGCTGCTTCCCACTGGCTT	1289		
QY	1344	CCTCATCTTCAATATCTTCTACTGGGTGTCTTATAAAGTGTCTATGTCAGAAATATCCA	1403		
DB	1290	CCTCATTTTCAACATCTTTTACTGGATCACATCAAGATCATTCGGCATGAAGATGTCCA	1349		
QY	1404	CCAG	1407		
DB	1350	CAAG	1353		

```

RESULT 11
RNI6RAA2
LOCUS      RNI6RAA2      3160 bp      mRNA      linear      ROD 23-MAR-1993
DEFINITION R.norvegicus mRNA for inhibitory glycine receptor alpha 2A subunit.
ACCESSION  X61159
VERSION     X61159.1  GI:288344
KEYWORDS   inhibitory glycine receptor alpha subunit.
SOURCE      Rattus norvegicus (Norway rat)
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE  1 (bases 1 to 3160)
AUTHORS   Kuhse,J., Kuryatov,A., Maulet,Y., Malosio,M.L., Schmieden,V. and
            Betz,H.
TITLE      Alternative splicing generates two isoforms of the alpha 2 subunit
            of the inhibitory glycine receptor
JOURNAL    FEBS Lett. 283 (1), 73-77 (1991)
MEDLINE    91243883
PUBMED     1645300
FEATURES   Location/Qualifiers
            source          1..3160
                        /organism="Rattus norvegicus"
                        /mol_type="mRNA"
                        /db_xref="taxon:10116"
                        532..1910
            CDS             /codon_start=1
                        /product="inhibitory glycine receptor alpha 2A subunit"
                        /protein_id="CAA43471.1"
                        /db_xref="GI:288345"
                        /db_xref="GOA:P22771"
                        /db_xref="UniProt/Swiss-Prot:P22771"
                        /translation="MNRQLNLTALFAFLGNHFRFAFKDHSRSKHPSQTLSP
                        SFLDKMRTSGYDRIEDIPNFKPEVNVTCNIFNSFGSVTETDNDKLLRISKNGKV
                        NDSRLAYSYPDSDLDSDMLDSIWKPLDFPANEKGNFHDVTDNDKLLRISKNGKV
                        LYSRLTLSCPMDLKNFPMQVOTCTMOLESFGVTMDNLIPEWLSGDPVQVAGELTL
                        POFILKEELGCTKHYNTGKFTCEVKFHLEROMGYLIOMYIPSLILVILSVSEF
                        WINMDAPARVALGITVLTWTQSSGRASLPKYSYKVAIDIWAVCLLFVPAALLE
                        YAAVNFVQKHFEFLRLRRQRQNKEDVTRESRNFSGYGMGHCLQVKGDTAVKAT
                        PANPLPQPPKQDAIKKFKVDRAKRIDTISRAAFPLAFILFNI FYWITYKIIRHEDVH
                        KK"
            sig_peptide      552..632
                        /note="inhibitory glycine receptor alpha 2A subunit"
            mat_peptide      631..1907
                        /product="inhibitory glycine receptor alpha 2A subunit"

ORIGIN
Query Match      39.9%; Score 654.8; DB 10; Length 3160;
Best Local Similarity 70.8%; Pred. No. 1.2e-168;
Matches 938; Conservative 0; Mismatches 332; Indels 54; Gaps 3;

QY 84 AAAGAGGAGTCAAAATCTGGAAACCAAGGGGTCCAGCCCATGTGCCCTCTGATTTCT 143
DB 635 AGACCATGACTCCAGGTCTGGAAACATCCCTCGCAGACCTGTCTCTCTCAGATTCTT 694
QY 144 AGCAAACTTATGGGGGACATCTCGATATGATGCCAGGATTCGGCCCAATTTAAAGG 203
DB 695 GGATAAATAATGGGAAGGACATCAGGATATGATGCAAGAATCAGGCCAAATTTAAAGG 754
QY 204 CCACCCCGTGAAGCTGACCTGCAACATCTTCATCAACAGTTTCAGCTCCGTCAACAGAC 263
DB 755 TCCTCCAGTAAGCTTACTTGCAATATTTTATCAACAGTTTGGATCGGTACAGAAAC 814
QY 264 CACAATGGACTACCGGGTGAATCTTCTTGGGGCAACAGTGAATGACCCAGCCTGTC 323
DB 815 CACCATGGACTACCGAGTGAACATTTTCTGAGACAGCAGTGGAAAGATTCACGGCTGCG 874
QY 324 CTACCGAGATATCCTGATGACTCTCTGGACCTCGATCCCTCCATCGTCTGACTCTATCTG 383
DB 875 ATACAGTGAATACCCAGATGATTCCTCGGATTTGGATCCATCGATGCTGATTCGATTTG 934

```

```

QY 384 GAAGCCAGACCTCTTCTTTGCTAATGAGAAAGGGGCCAACTTCCATGAGGTGACCAACGGA 443
DB 935 GAAACAGATTGTTCTTTGCTCAATGAGAAAGAGGCCAATTTCCATGATGTCAACACTGA 994
QY 444 CAACAGTTACTGCGCATCTTCAAGATGGGAATGTGCTGTACAGCATCAGGCTGACCTT 503
DB 995 TAAACAAGTTGCTGCGGATTTCCAAAAATGGCAAGTGCTCTACAGTATTAGACTCACCTT 1054
QY 504 CATTTTGTCTGCTGCTGAGTGGACCTCAAGAACTTCCCATGGACATCCAGACCTGGACGAT 563
DB 1055 GACTTTATCTGCTCCCATGGACCTGAAGAACTTTTCAATGGATGTCCAGACCTGTACAAT 1114
QY 564 GCAGCTTGAGAGCTCATCCATCTCTGCAGCCCTCTGCCATCTCTGTCTACTTTCAGTTGG 623
DB 1115 GCAGCTGGAGATT-----TTGG 1132
QY 624 CTACACCATGAAAGACCTGCTGTTTGAAGTGGCTGGAAGATGCTCTCTGTCTCAAGTGGC 683
DB 1133 GTACACCATGAATGACCTGATATTTGAGTGGTTAAGTGTATGGTC---CAGTACAAAGTTGC 1189
QY 684 TGAGGGGCTGACTCTGCCCTGCTTATCTTGGCGGATGAGAGGATCTAGGCTGTGTAC 743
DB 1190 TGAAGGACTCACCTGCTGCTCAGTTTATTTTGAAGAGAGAGAAAGAACTTTGGCTATTGGCAC 1249
QY 744 CAAGCACTACAACACAGAGGAAATTCACCTGCAATCGAGGTAAAGTTTCACTCGAAACGCA 803
DB 1250 AAAGCATTAACACACTGGCAAGTTTACCTGCAATGAGGTCAAGTTTCACTCGAAGCCCA 1309
QY 804 GATGGGCTACTATCTGATTCAGATGTACATCCAGCCTTACTCATCGTGTCTCTCTGCTG 863
DB 1310 GATGGGCTATTATTTGATCCAGATGTATATCCAGCCTTACTCATGATGATCATTTTGTCTG 1369
QY 864 GGTCTCTCTGATCAACATGATGCTGCCCTGCCGTGTGGCGCTGGGCTGACCAAC 923
DB 1370 GGTCTCTCTTGGATAAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1429
QY 924 CCGTCTCAACATGACACACACAGAGCTCTGCTGCTCCCGGGCTCTTTCGCTTAAAGTGTCTTA 983
DB 1430 CGTCTGCAATGACTACCCAGAGTTTCAAGTTTCCAGGGCATCTCTGCCAAAGTCTCATA 1489
QY 984 CCGTGAAGCAATCGACATCTGATGGCTGTGTGTCTGTCTTCTTGTGTGTGTGTGTGTGTGT 1043
DB 1490 TGTGAAGCAATGACATCTGATGGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1549
QY 1044 CGAGTATGCTGCATATAAATTTTGTCTGTCTGATGATGATGATGATGATGATGATGATGAT 1103
DB 1550 GGAATATGAGCAGTGAACCTTGTCTCCAGGCAACACAGAGGATTTCTTCGTCTCCGGAG 1609
QY 1104 AAAGCAGAGCGCCCAACCGCTTGGAGGAAGATATCATCAAGAAAGTCGTTTCTATTTCGG 1163
DB 1610 ACGACAGAAGAGCGCAGATAAGAGAGATGTTACTCGTGAAGCCGTTTAACTTCAAG 1669
QY 1164 TGGCTATGCTTGGGCGCATCTGCTGAGCAAGATGAGGATCCATGGAAGGTTTCTGG 1223
DB 1670 CGGTTATGGGATGGGTCTGCTGCTCCAAAGTGAAGATGATGATGATGATGATGATGATGATGAT 1726
QY 1224 CATTTATAGTCCCAACCTCCAGCCCTTCTTAAAGGAAGGAGAAACCAACCGGAAACT 1283
DB 1727 -----ACCTGCCAACCCACTTCCGCAACCCCAAAAGATGAGATGCTATCAAGAGAA 1780
QY 1284 CTACGTGGAAGAGCAAGAGAAATGACACCATCTCCCGGGCTGTCTTCCCTTTTCACTTT 1343
DB 1781 GTTTGTGGATCGGCAAAAGAAATGACACCATATCTCGAGCTGCTTCCCACTGSCCTT 1840
QY 1344 CTTCTATCTTCAATATCTTCTACTGGGTTGTCTATTAAGTGTCTATGGTCTGAGAGATATCCA 1403
DB 1841 CCTCATTTTCAACATCTTTTACTGGATCACATACAGATCATTCGGCATGAAGATGTCCA 1900
QY 1404 CCAG 1407
DB 1901 CAAG 1904

```

```
RESULT 12
AX700251
LOCUS AX700251 3865 bp DNA linear PAT 03-APR-2003
DEFINITION Sequence 20 from Patent EP1284297.
ACCESSION AX700251
VERSION AX700251.1 GI:29536131
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1
AUTHORS Brooksbank, R.A., Dixon, A.K., Lee, K. and Pinnock, R.D.
TITLE Identification and use of molecules implicated in pain
JOURNAL Patent: EP 1284297-A 20 19-FEB-2003;
WARNER-LAMBERT COMPANY (US)
FEATURES
source
1..3865
/organism="Rattus norvegicus"
/mol_type="unassigned DNA"
/db_xref="taxon:10116"
/notes="Neonatal glycine receptor"
ORIGIN
Query Match 39.9%; Score 654.8; DB 6; Length 3865;
Best Local Similarity 70.8%; Pred. No. 1.2e-168;
Matches 938; Conservative 0; Mismatches 332; Indels 54; Gaps 3;
QY 84 AAAGAGGAGTCAAAATCTGGACCAAGGGTCCAGCCCATGTCCTCCCTCTGATTTCT 143
DB 1256 AGACCATGACTCCAGTCTGGAAAACATCCCTCGACAGCCTGTCTCTTCAGATTCTT 1315
QY 144 AGCAAACTTATGGGCGAACATCTGGATATGATGCCAGGATTCGCCCAATTTTAAAG 203
DB 1316 GGATAAATAATGGGAAGGACATCAGATATGATGCAAGATTCAGGCCAAATTTTAAAG 1375
QY 204 CCACCCGTGAAGTCACTGCAACATCTTCATCAACAGTTTCAGTCCGTTCACCAAGAC 263
DB 1376 TCCCTCAGTAAAGCTTACTTGCATATTTTATCAACAGTTTGGATCGGTTCACAGAAC 1435
QY 264 CACAATGGACTACCGGTGAATCTCTTCTGGGCAACAGTGAATGACCCAGCCTGTC 323
DB 1436 CACCATGGACTACCGAGTGAACATTTCTGAGACAGCAGTGAAGATTCAGCGTGGC 1495
QY 324 CTACCGAGATATCTGTGACTCTCTGGACCTCGATCCCTCCATCGCTGAGACTCTATCTG 383
DB 1496 ATACAGTGAATCCAGATGATTCCTCGATTTGGATCCATCGATCGTGGATTCGATTG 1555
QY 384 GAAGCAGACCTCTCTTTGCTAATGAGAAAGGGGCAATCTTCATGAGTGACACCGA 443
DB 1556 GAAACCAAGATTTGTTCTTTGCCAATGAGAAAGGAGCAATTTCCATGATGTCCACACTGA 1615
QY 444 CAACAGTGTACTCGCATCTCAAGATGGATGTGCTGACAGCATCAGGCTGACCT 503
DB 1616 TACAAGTGTCTCGGATTTCCAAAATGGCAAATGCTCTACAGATTAGACTACCTT 1675
QY 504 CATTTGTCTGCTGATGACCTCAAGAACTTCCCATGGAATCCAGACGTGACAGAT 563
DB 1676 GACTTTATCTGTCCTGACCTGAGAACTTTCCAAATGGATGTCCAGACCTGTACAT 1735
QY 564 GCAGCTTGAGAGCTATCCATATCTCTGACGACCTCTGCGCATCTCTGTCACTTTCA 623
DB 1736 GCAGCTGGAGAGTT-----TTGG 1753
QY 624 CTACACCATGAAGACTCGTGTTCAGTGGCTGGAAGATGCTCTGCTGTCCAGTGGC 683
DB 1754 GTACACCATGAATGACTGATATTTTGGTGGTTAAAGTATGGTCT---CAGTACAGTTGC 1810
QY 684 TGAGGGGCTGACTCTGCCCCAGTTTATCTTGGGGATGAGAAGCATCAGGCTGTGTGATC 743
DB 1811 TGAAGACTCACCTGCTCAGTTTATTTTGAAGAAGAGAGAACTGGCTATTGGCAC 1870
```

```
744 CAAGCACTACACACAGGGGAATTCACCTGCTCGAGGTAAAGTTTCACTCGAAGCGCA 803
1871 AAAGCATTAACAACACTGGCAAGTTTACCTGCAATTTAGGTCAAGTTTCACTCGAAGCGCA 1930
804 GATGGCTACTATCTGATTCAGATGATACATCCCGAGCTTACTCATCGTCACTCTGCTCTG 863
1931 GATGGCTATTATTGATCCAGATGATATCCCGAGCTTCTCATAGTCACTTTGCTCTG 1990
864 GGTCTCTCTCGATCAACATGGATGCTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 923
1991 GGTCTCTCTTTGGATAAACATGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2050
924 CGTGCTCACCATGACCCAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 983
2051 CGTGCTGACAATGACTTACCAGAGTTTACAGTTCAGGGTTCAGGGCATCTCTGCCAAGAGTCTCA 2110
984 CGTGAAGGCAATCGACATCTGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1043
2111 TGTGAAGCAATGACATCTGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2170
1044 GGAGTATGCTGCATAAATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1103
2171 GGAATATGACGAGTGAATTTGTCTCCAGGCAACACCAAGGAGTTCCTTCTGCTGCTGCTGCTG 2230
1104 AAGGAGAGCGCCCAACGCTTGGAGAGATATCATCCAGAAAGTCTGTTCTATTTCG 1163
2231 ACACAGAGAGGCGCAGATAAAGAAAGATGTACTGCTGAAAGCGTTTAACTTCAG 2290
1164 TGGCTATGCTTGGGCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1223
2291 CGTTATGGATGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2347
1224 CATTTATGCTCCCAACCTTCCAGCCCTCTTTAAGGGAAGGAGAAACCAACGCGGAACT 1283
2348 -----ACCTGCCAACCCACTTCCGCAACCCCAAAAGATGCAAGATGCTATCAAGAAGAA 2401
1284 CTAGTGGACTGACCAAGAGATTTGACACCATCTCCCGGCTGCTTCCCTTTCACTTT 1343
2402 GTTGTGGATGGGCAAAAGAAATCGACCATATCTCGAGCTGCTTCCACATGCGCTT 2461
1344 CTTCTATTTCAATATCTTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1403
2462 CTTCTATTTCAATATCTTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2521
1404 CCAG 1407
2522 CAAG 2525
```

```
RESULT 13
RNNEOGLY
LOCUS Rat NGLYr mRNA for neonatal glycine receptor.
DEFINITION Rat NGLYr mRNA for neonatal glycine receptor.
ACCESSION X57281
VERSION X57281.1 GI:56743
KEYWORDS glycine receptor; NGLYr gene.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 3865)
AUTHORS Akagi, H., Hirai, K. and Hishinuma, F.
TITLE Cloning of a glycine receptor subtype expressed in rat brain and
spinal cord during a specific period of neuronal development
JOURNAL FEBS Lett. 281 (1-2), 160-166 (1991)
, MEDLINE 91200276
PUBMED 1707830
REFERENCE 2 (bases 1 to 3865)
AUTHORS Hishinuma, F.
TITLE Direct Submission
JOURNAL Submitted (28-JAN-1991) F. Hishinuma, Mitsubishi Kasei Institute of
Life Sciences, 11 Minamiooya, Machida-shi, Tokyo 194, Japan
```

FEATURES	Location/Qualifiers	Db	1754	GTACACCAATGAATGACCTGATATTTAGTGGTTAAGTGTATGGTC---CAGTACAAGTTCG	1810		
source	1. .3865 /organism="Rattus norvegicus" /mol_type="mRNA" /strain="Wistar" /db_xref="taxon:10116" /tissue_type="spinal cord" /dev_stage="10 days old" 1. .3865 /gene="NGLYR" 1. .3865 /gene="NGLYR" /evidence=experimental 1173. .2531 /gene="NGLYR" /codon_start=1 /product="neonatal type of glycine receptor" /protein_id="CAA40549.1" /db_xref="GI:56744" /db_xref="UniProt/Swiss-Prot:P22771" /translation="MNQVNLITLAFELGTHREAFKDHRSKGKPSQTLSP SDFDLKMGRTSGYDRIAPNFKPVNTCNIPINSFGSVETETMDYRVNLFLEQOW NLSRLAYSEPPDSDLDPNLSIDIWKPDLFANEGKGNFHDVTDNKLRLSKNGKV LYSIRLTLSCPMDLNFMNDVQCTMQLFESFGYTMNDLIFEWLSDFGVQVAGLTL POFLKEELGCTKHYNTKFTCEVFKHLEROMGYLIQWIPSLILVILSNWSP WINMDAAPARVALGITTTLTMTTQSSGRSLPKSVYKVIDIIMWAVCLLFFAALLE YAAVNFVSQHKLEFLRLRRQKNEEDVTRESRNFSGYGNHCLQV/KDGTAVKAT PANPLPQPPADADAIIKKFVDRNRIDTISRAPPFLAFLIFNYITTYIKIRHEDVH KK"	Qy	684	TCAGGGGCTGACTCTGCCCCAGTTTATCTTGGGATGAGAAGGATCTAGGCTGTGTAC	743		
		Db	1811	TGAAGGACTCACCTGCTCAGTTTATTTTGAAGAAGAGAGAAAGAACTTGGCTATTGCAC	1870		
gene		Qy	744	CAAGCACTACAACACAGAGGAAATTCACCTGCATCGAGGTAAAGTTTCACTCGAAGCGCA	803		
mRNA		Db	1871	AAAGCATTAACAACACTGGCAAGTTTACCTGCTTACCTGCTTACCTGGAAGCGCA	1930		
		Qy	804	GATGGGCTACTATCTGATTCAGATGATACATCCAGCCTTACTCATCGTTCATCCTGTCCTG	863		
CDS		Db	1931	GATGGGCTATTATTGATCCAGATGATATCCAGCCTTCCAGCCTGCTGATAGTCAITTTGTCTG	1990		
		Qy	864	GCTCTCTCTCTGATCAACATGCTGCCCTGCCCTGCTGGCTGGGCTGGGATCACCAAC	923		
		Db	1991	GCTCTCTCTCTGATCAACATGCTGCCCTGCCCTGCTGGCTGGGATCACCAAC	2050		
		Qy	924	CCTGCTCACCATGACACCCAGAGCTCTGGCTCCCGGGCTCTTTGCTTAAGTGTCTCTA	983		
		Db	2051	CCTCTGACAAATGACTACCCAGAGTTTCAGGTTCCAGGCGATCTCTGCCAAGGTCTCTA	2110		
		Qy	984	CCTGAAGCAATCGACATCTGGATGGCTGTGTCTGCTCTTTGTGTTGCTGCTGCTTGTCT	1043		
		Db	2111	TGTGAAGCAATGACATCTGGATGGCAGTGTGCTTCTGTTGTGTGTTGCTGCTGCTTACT	2170		
		Qy	1044	GGAGTATGCTGCCATAAATTTGTTCTGTCAGCATAAAGATTCATACGACTTCGAAG	1103		
		Db	2171	GGAATATGACAGTGAACCTTTGCTCCAGGCAACACAAAGGAGTTCCTTCGCTCCGGAG	2230		
ORIGIN		Qy	1104	AAGGACAGAGCGCCAAACGCTTGGAGAGATATCATCAAGAAAGTCGTTTCTATTTCGG	1163		
	Query Match 39.9%; Score 654.8; DB 10; Length 3865; Best Local Similarity 70.8%; Pred. No. 1.2e-168; Matches 938; Conservative 0; Mismatches 332; Indels 54; Gaps 3;	Db	2231	ACGACAGAGAGCGCAGATAAGAGAGATGTTACTCTGGAAGCGTTTAACTTCAG	2290		
		Qy	1164	TGGCTATGGCTTTGGGCCACTGCTGCGAGCAAGATGAGGTCCAAATGGAAGTTCTCG	1223		
		Db	2291	CGGTTATGGGATGGGTCACTGCTCCAAAGTGAAGATGGTACAGCTGTCAAGGCTAC---	2347		
		Qy	1224	CATTTATAGTCCCAACCTCCAGCCCTCTTCTAAGGAGAGGAGAAACCCAGCGGAACT	1283		
		Db	2348	-----ACCTGCCAACCCACTTCCGCAACCCCAAAAGATGAGATGCTATCAAGAAGAA	2401		
		Qy	1284	CTACGTGGACTGAGCAAGAGAAATGACACCATCTCCCGGGCTGTCTTCCCTTTTCACTTT	1343		
		Db	2402	GTTTGGATCGGCAAAAGAAATCGACACCATATCTCGAGCTGCTTCCCACTGSCCT	2461		
		Qy	1344	CCTCATCTTCAATATCTTCTACTGGTTGTCTATAAAGTGTATGTCAGAGATATCCA	1403		
		Db	2462	CCTCATTTTCAACATCTTTTACTGGATCACAACAGATCATTCGGCATGAAGATGTCCA	2521		
		Qy	1404	CCAG 1407			
		Db	2522	CAAG 2525			
		RESULT 14	RNO310834	1350 bp	mRNA	linear	ROD 17-MAR-2004
		LOCUS	Rattus norvegicus	1350 bp	mRNA	linear	
		DEFINITION	Rattus norvegicus	1350 bp	mRNA	linear	
		ACCESSION	AJ310834	1350 bp	mRNA	linear	
		VERSION	AJ310834.1	1350 bp	mRNA	linear	
		KEYWORDS	alternative splicing; glycine receptor alpha 1 precursor.				
		SOURCE	Rattus norvegicus (Norway rat)				
		ORGANISM	Rattus norvegicus				
		REFERENCE	1				
		AUTHORS	Groot-Kormelink, P.J., Beato, P.J., Finotti, C., Harvey, R.J. and Sivilotti, L.G.				
		TITLE	Achieving optimal expression for single channel recording: a				

ORIGIN

Query Match	39.9%;	Score 654.8;	DB 10;	Length 3865;
Best Local Similarity	70.8%;	Pred. No. 1.2e-168;		
Matches 938;	Conservative 0;	Mismatches 332;	Indels 54;	Gaps 3;
Qy	84	AAAGAGGAGTCAAACTCTGGAAACCAAGGGGTCCAGCCCATGTCCTCCCTCTGATTTCCT	143	
Db	1256	AGACATGACTCCAGGTCTGGAAACATCCCTCGCAGACCTGTCTCTTCAGATTTCCT	1315	
Qy	144	AGACAACTTATGGGGCAACATCTGGATATGATGCCAGGATTCGGCCCAATTTTAAAGG	203	
Db	1316	GGATAAATAATGGGAAGGACATCAGGATATGATGCAAGATCAGGCCAAATTTTAAAGG	1375	
Qy	204	CCACCGGTGAAGTCAAGTCAACATCTTCAACAGTTTTCAGCTCGTCCAGCAAGAC	263	
Db	1376	TCCTCAGTAAAGCTTACTTGGCAATTTTATCAACAGTTTGGATCGGTCAAGAAAC	1435	
Qy	264	CACAATGGACTACCGGTGAATGCTCTTTCGGCGCAACAGTGAATGACCCAGCCCTGTC	323	
Db	1436	CACCATGGACTACCGAGTGAACATTTTCTGAGACAGCAGTGAACGATTCACGGCTGGC	1495	
Qy	324	CTACCGAATATCTGATGACTCTCTGGACCTCGATCCCTCCATGCTGGACTCTATCTG	383	
Db	1496	ATACAGTGAATACCGAGATGATTCCTTGGATTTGGATCCATGATGCTGATTCGATTTG	1555	
Qy	384	GAAGCCAGACTCTCTTTCTGATGAGAAAGGGCCAACTTCATGAGTGAACACCGGA	443	
Db	1556	GAACACAGATTTGTTCTTTGGCAATGAGAAAGGAGCAATTTCCATGATGTACCACTGA	1615	
Qy	444	CAACAAGTTACTCGCATCTTCAAGAAATGGAAATGTCTGTACAGCATCAGGCTGACCT	503	
Db	1616	TAACAAGTTGCTCGGATTTCCAAAATGGCAAAAGTCTCTACAGTATTAGACTCACTT	1675	
Qy	504	CATTTTGTCTGCTGATGAGCACTCAAGAACTTCCCATGGACATCCAGACGTGCACGAT	563	
Db	1676	GACTTTATCTCTCCATGACCTGAAGAACTTTTCCAAATGGATGTCCAGACCTGTACAAT	1735	
Qy	564	GCAGCTTGAGAGTCAATCCATCTGTCAGCCCTCTGCCATCTCTGTCACTTTCAGTTGG	623	
Db	1736	GCAGCTGGAGAGTT-----TTGG	1753	
Qy	624	CTACACCATGAAGACCTCGTGTGTTGAGTGGTGGAGATGCTCTCTGCTGCCAAGTGC	683	

QY	1104	AAGGCAGAGGCGCCAAAGCTTTGGAGGAATATATCACAAGAAAGTCTGTTCTATATTCG	118	
Db	2231	ACGACAGAGAGGCGAGATAATAGGAAGAAGATGTTTACTCGTGAAGACCGTTTTAACTTCAG	2290	
QY	1164	TGGCTATGCTTTGGGCCACTGCCTCGAGGCAAGAGATGGAGGTCCAATGGAAGTTCTCG	1223	
Db	2291	CGTTATGGATGGGTCTACTGCCCTCCAGTGAAGATGGTACAGCTGTCAAGGCTAC---	2347	
QY	1224	CATTTATAGTCCCAACCTCCAGCCCTCTTCTAAGGGAAAGGAGAAACACGCGGAAACT	1283	
Db	2348	-----ACCTGCCAACCCACTTCCGCAACCCCAAAAAGATGCAGATGCTATCAAGAAGAA	2401	
QY	1284	CTACGTGACCTGAGCAAGAGAATTGACACCATCTCCGGGGCTGTCTTCCCTTTCACTTT	1343	
Db	2402	GTTTGTGGATCGGGCAAAAAGATCGACACCATATCTCGAGCTGCCTTCCCACTGGCCCT	2461	
QY	1344	CCTCATCTTCAATATCTTCTACTGGGTGTCTATAAAGTGTCTATGGTGCAGAGATATCCA	1403	
Db	2462	CCTCATTTTCAACATCTTTTACTGGATCACATACAAGATCATTCGGCATGAAGATGCCA	2521	
QY	1404	CCAG 1407		
Db	2522	CAAG 2525		
RESULT 14				
RNO310834				
LOCUS				
DEFINITION				
	RNO310834	1350 bp	mRNA	linear
				ROD 17-MAR-2004
				Rattus norvegicus mRNA for glycine receptor alpha 1 precursor,
				primary transcript.
ACCESSION				
	AJ310834			
VERSION				
	AJ310834.1	GI:13548654		
KEYWORDS				
				alternative splicing; glycine receptor alpha 1 precursor.
SOURCE				
				Rattus norvegicus (Norway rat)
ORGANISM				
				Rattus norvegicus
				Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
				Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
				Rattus.
REFERENCE				
	1			
AUTHORS				
				Groot-Kormelink,P.J., Beato,P.J., Finotti,C., Harvey,R.J. and
				Sivilotti,L.G.
TITLE				
				Achieving optimal expression for single channel recording: a

plasmid ratio approach to the expression of alphanal glycine  
 receptors in HEK293 cells  
 J. Neurosci. Methods 113, 204-214 (2002)

2 Beato, M., Groot-Kormelink, P.J., Colquhoun, D. and Sivillotti, L.G.  
 Openings of the rat recombinant alpha 1 homomeric glycine receptor  
 as a function of the number of agonist molecules bound  
 J. Gen. Physiol. 119 (5), 443-466 (2002)

J. Neurosci. 21977760  
 11981023

3 Burzomato, V., Groot-Kormelink, P.J., Sivillotti, L.G. and Beato, M.  
 Stoichiometry of recombinant heteromeric glycine receptors revealed  
 by a pore-lining region point mutation  
 Recept. Channels 9 (6), 353-361 (2003)

14698963

4 Beato, M., Groot-Kormelink, P.J., Colquhoun, D. and Sivillotti, L.G.  
 The activation mechanism of alphanal homomeric glycine receptors  
 J. Neurosci. 24 (4), 895-906 (2004)

14749434

5 (bases 1 to 1350)  
 /organism="Rattus norvegicus"  
 /mol\_type="mRNA"  
 /strain="Sprague-Dawley males"  
 /db\_xref="taxon:10116"  
 /tissue\_type="whole brain"  
 /clone\_lib="rat brain marathon-ready cDNA"  
 /dev\_stage="8-12 weeks"

1. .1350  
 /function="ion channel"  
 /notes="alternative splicing, primary transcript"  
 /codon\_start=1  
 /evidence=experimental  
 /product="glycine receptor alpha 1 precursor"  
 /protein\_id="CAC35978.1"  
 /db\_xref="GI:13548655"  
 /translation="MYSFNTLRFLMETIVFFSLAASKEADAARSAPKPMSPDFLKL  
 LMGRTSGYARIRPNKGPVNVSCNIFINSGFIAETTMVYRVNIFLRQWMDRLA  
 VNEYPDSDLDPSMLDSIKPDLPEANEKGFHEITTDNKLRLSRNGVLYSIRI  
 TITLACPMOLKHPMDVQICIMQLESFGYTMNDLFEWQEGAVQVADGLTPQFLK  
 EKLDRYCTKHYNTGFTCEARFHLERQMYLIQMYIPSLIIVLSWISFWINMDA  
 APARVLGITTTLTMTQSSGRASLPKVSYYKVIDIMWAVCLLFVFSALLLYAAVNF  
 VSRHKELLRFRKRHRHDDGEGRFNFSAYMGPAQLQAKDGISVKGANNNTTN  
 PAPAPSKSPFEMREKFLIQRAKKIDKISRIGFPMAPFLIFNMFYIIVREDVHNK"

CD5

Query Match 39.9%; Score 654.6; DB 10; Length 1350;  
 Best Local Similarity 72.6%; Pred. No. 1.2e-168;  
 Matches 941; Conservative 0; Mismatches 299; Indels 57; Gaps 5;

QY 116 CCAGGCCATGTCCTGATTTCTAGACAACTATGGGGCAACATCTGGATATG 175  
 |||||  
 DB 98 CCAAGCCTATGTCACCTCGGACTTCTCGATAAGCTTATGGGAAGACTTCTGGGTATG 157  
 |||||

QY 176 ATGCCAGGATTCGGCCCAATTTTAAAGGCCACCCGCTGAACGTCGCTGCAACATCTTCA 235  
 |||||

DB 158 ATGCCAGGATTCAGACCCCAACTTTAAAGGTCCTCTGTGAACGTCGAGTTCGAACATCTTCA 217  
 |||||

QY 236 TCAACAGTTTTCAGCTCCGTCACCAAGACCAACATGGACTACCGGGTGAATGCTCTTTCG 295  
 |||||

DB 218 TCAACAGCTTTGGTTCTATCGCGGAGACCAACCATGACTACAGGGTCAACATCTTCTCTGA 277  
 |||||

QY 296 GGCAACAGTGGAAATGACCAACGCTGTCTTACGAGAATATCTGATGACTCTCTGGACC 355  
 |||||

278 GGCAGCAGTGGAAACGACCCCGCTCTCGCTACATGAATATCTCTGACGACTCTCTGGACC 337  
 QY

356 TCGATCCCTCCATGCTGGACTCTATCTGGAAGCCAGACCTCTTCTTTGTAAATAGAAAG 415  
 |||||

DB 338 TTGACCCATCCATGTTGGATTCCATCTGGAAGCCTGACTTGTCTTTGCCAATGAGAAG 397  
 |||||

QY 416 GGGCCAACTTCCATGAGGTGACACGGCAACAAGTTACTGCGCATCTTCAAGAAATGGGA 475  
 |||||

DB 398 GGGCCCACTTCCACGAGATCACACGGCAACAAGCTGCTGAGAACTCTCCCGAAACGGCA 457  
 |||||

QY 476 ATGTGCTGTACAGCATCAGGCTGACCTCATTTTGTCTGCTGCTGATGGACCTCAAGAACT 535  
 |||||

DB 458 AGTCTCTACAGCATCAGAACTACCCCTGACTCTGGCTGCCCATGGACCTCAAGAAAT 517  
 |||||

QY 536 TCCCATGGAATCAGACGTCGACGTCGACGTCGAGAGTCATCCATCTCTGACGCC 595  
 |||||

DB 518 TCCCATGGAATCAGACGTCGACGTCGACGTCGAGAGTCATCCATCTCTGACGCC 559  
 |||||

QY 596 CTCTGCCATCTCTGTCACCTTTCAGTTGGCTACACCATGAAAGACCTCGTTTGGATGGC 655  
 |||||

DB 560 -----TTGGTTATACCATGAACGACCTCATCTTTTGGATGGC 595  
 |||||

QY 656 TGGAAAGATCTCTCTGCTGTCACAGTGGCTGAGGGGCTGACTCTGCCCCAGTTTATCTTGC 715  
 |||||

DB 596 ----NAGACAAAGAGCTGTGAGGTGGCAGACGACTGACCTGCTCAGTTTATCTGA 652  
 |||||

QY 716 GGGATGAGAAGGATCTAGGCTGTGTGACCAAGCAGCTACAAACAGGGAAATTCACCTGCA 775  
 |||||

DB 653 AGGAAGAGAAGGATCTGAGATATGTCACCAAGCAGCTACAAACAGGTAATTCACCTGCA 712  
 |||||

QY 776 TCGAGGTAAGTTTACCTGGAACGGCAGATGGGCTACTATCTGATTTCAGATGTATATCC 835  
 |||||

DB 713 TTGAGGCCCGATTCACCTGGAAACGGCAGATGGGCTACTACCTGATCCAGATGTATATCC 772  
 |||||

QY 836 CCAGCTACTCTCATCTGTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 895  
 |||||

DB 773 CCAGCTGCTTATGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 832  
 |||||

QY 896 CTGCCCCGTGTCGGCTGCGCATCACACCGTCTCACCATGACCAACCCAGAGCTCTGGCT 955  
 |||||

DB 833 CAGCTGCTGTCGGCATCTGGGCATCACACAGTGTCTACCATGACCAACACAGAGCTCTGGCT 892  
 |||||

QY 956 CCGGGCTCTTTGCTTAAAGTGTCTGCTGTAAGGCAATTCACATCTCGATGGCTGTGT 1015  
 |||||

DB 893 CCGAGCCTCTCCATCCCAAGGTCCTCTGCTGTAAGGCTTATGACATCTGATGGCTGTTT 952  
 |||||

QY 1016 GTCTGCTCTTCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1075  
 |||||

DB 953 GCTGCTCTTCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1012  
 |||||

QY 1076 AGCATAAAGAAATTCATACGACTTTCGAAGAGGACGAGCGGCCAACGCTTTGGAGGAAGATA 1135  
 |||||

DB 1013 AACACAAAGAACTC-----CTTCGATTTAGGAGGAAGCGGACATCAAGGATGATG 1066  
 |||||

QY 1136 TCATCAAGAAAGTGTCTTCTATTTTCGTGGCTATGCTTGGGCTC---ACTGCTCGAGG 1192  
 |||||

DB 1067 AGGGTGGAGAGCGCGCTTCAACTTCTCCGCTCTGAGGATGGGCGCCAGCTGCTGCTCAAG 1126  
 |||||

QY 1193 CAAAGATGGAGGTCCAATGGAAGTTCCTGCGATTTATAGTCCCCAACCTCTCCAGCCCT- 1251  
 |||||

DB 1127 CCAAGATGGCATCTCTGTCAAGGGTGCCAAACAAACACACACGACGACCCCGCTCTG 1186  
 |||||

QY 1252 --CTTCTAAGGAAAGGAGAAACACCGCGAAACTCTACCTGGAGCTGAGCAAGAGAAATG 1309  
 |||||

DB 1187 CACCGTCCAAGTCCCGGAGGAGATGCGGAACCTCTTCATCCAGAGACCAAGAGATCG 1246  
 |||||

QY 1310 ACACATCTCCCGGGGTGCTTCCCTTTCATCTTCTCTCATCTTCAATATCTTCTACTGGG 1369  
 |||||

DB 1247 ACAAATAATCTCGCATCGGTTTCCCAATGGCTTCTCTCATCTTCAATGATGTTCTACTGGA 1306  
 |||||

QY 1370 TTGTCTATAAAGTGTATGTTGTCAGAGATATCCACA 1406  
 |||||

DB 1307 TCATCTACAAGATCGTTCGGAGAGAGGACGTCACAA 1343  
 |||||



Qy 1370 TTGCTATAAGTCCTATGTCAGAGATATCCACCA 1406  
Db 1319 TCATCTATAAGATCGTCCGGAGAGAGGATGTCCCAA 1355

Search completed: April 28, 2005, 14:18:14  
Job time : 7725.53 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 28, 2005, 01:49:33 ; Search time 1076.49 Seconds  
(without alignments)  
9018.523 Million cell updates/sec

Title: US-10-075-846-3  
Perfect score: 1640  
Sequence: 1 atgacaactctgttcctgc.....taaaaaaaaaaaaaaaaaaaaa 1640

Scoring table: IDENTITY\_NUC  
Gapex 10.0, Gapex 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*

- 1: Geneseq1980s:\*
- 2: Geneseq1990s:\*
- 3: Geneseq2000s:\*
- 4: Geneseq2001s:\*
- 5: Geneseq2001bs:\*
- 6: Geneseq2002as:\*
- 7: Geneseq2002bs:\*
- 8: Geneseq2003as:\*
- 9: Geneseq2003bs:\*
- 10: Geneseq2003cs:\*
- 11: Geneseq2003ds:\*
- 12: Geneseq2004as:\*
- 13: Geneseq2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1640	100.0	1640	6	AAL49660 Human HGR
2	1524.8	93.0	2565	6	AAL49659 Human gly
3	1521.2	92.8	1874	6	AAD33667 Human TRI
4	1503.8	91.7	1600	3	AAC61678 cDNA sequ
5	1504.6	61.3	1574	12	ADJ27240 Human TRI
6	947.8	57.8	1224	12	ADJ27254 Human TRI
7	817.8	49.9	1176	8	ACC46177 Human dit
8	807.8	49.3	993	6	AAL49665 Human gly
9	660.8	40.3	1746	10	ABV76108 Neuron-Sp
10	656	40.0	1715	13	ADQ89057 Human uro
11	654.8	39.9	3865	10	ADB85139 Rat neona
12	650.2	39.6	1857	8	ABZ18534 Group III
13	648.6	39.5	1386	10	ABA00851 GLRA rela
14	648.6	39.5	1854	10	ABA00850 GLRA rela
15	552.8	33.7	3069	10	ADC77686 Human 238
16	552.8	33.7	3069	10	ABA00854 GLRA rela
17	552.8	33.7	3069	13	ADQ89103 Human uro
18	551.8	33.6	2225	12	ADJ27249 Human TRI
19	521.4	31.8	2569	10	ABA00852 GLRA rela
20	520.8	31.8	1438	10	ABA00855 GLRA rela

21	479	29.2	679	6	AAL49661	Aal49661 Human gly
22	366.2	22.3	657	2	AAQ25722	Aaq25722 GABA-A re
23	299	18.2	554	4	AAD17170	Aad17170 Human ion
24	299	18.2	554	8	ACD01555	AcD01555 cDNA clon
25	299	18.2	554	10	ADJ29246	AdJ29246 Novel hum
26	288.8	17.6	1129	10	ABA00853	AbA00853 GLRA rela
27	253	15.4	253	10	ACA55673	AcA55673 Human sig
28	233	15.4	253	12	ADI55469	AdI55469 Human pol
29	232.6	14.2	2106	13	ADR24668	AdR24668 Breast ca
30	227.8	13.9	39796	3	AAC61681	Aac61681 Nucleotid
31	221	13.5	543	12	ACH76139	Ach76139 Human gen
32	221	13.5	2404	6	ABI99254	AbI99254 Mouse isc
33	219	13.4	492	9	ACH28631	Ach28631 Human adu
34	208	12.7	214	12	ACH89839	Ach89839 Human gen
35	204.4	12.5	1614	6	AAD22072	Aad22072 Dermacent
36	204.4	12.5	1614	6	AAD22070	Aad22070 Dermacent
37	202.8	12.4	1614	6	AAD22071	Aad22071 Dermacent
38	193.8	11.8	1628	10	ADP81662	AdP81662 Leukaemia
39	193	11.8	1555	2	AAT28989	Aat28989 Human GAB
40	191.4	11.7	1359	10	ADB78711	Adb78711 Human GAB
41	191.4	11.7	1359	10	ADB78693	Adb78693 Human GAB
42	191.4	11.7	1359	10	ADB78697	Adb78697 Human GAB
43	189.8	11.6	1359	6	ABK27334	Abk27334 Mutant ga
44	189.8	11.6	1359	10	ADB78712	Adb78712 Human GAB
45	189.8	11.6	1359	10	ADB78721	Adb78721 Human GAB

## ALIGNMENTS

RESULT 1  
AAL49660  
ID AAL49660 standard; cDNA; 1640 BP.  
XX  
AC AAL49660;  
XX  
DT 27-NOV-2002 (first entry)  
XX  
DE Human HGRA4 splice variant coding sequence SEQ ID NO: 3.  
XX  
KW Human; glycine receptor alpha subunit 4; HGRA4; HGRA4sv; splice variant;  
KW cardiovascular disorder; reproductive disorder; neural disorder;  
KW cardiac; antiarrhythmic; antianginal; antiarrheic; antiulcer;  
KW neurotropic; neuroprotective; antibacterial; virucide; protozoicide;  
KW nervous system disorder; gastrointestinal disorder; gene therapy;  
KW infection; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200266606-A2.  
XX  
PD 29-AUG-2002.  
XX  
PF 13-FEB-2002; 2002WO-US0004329.  
XX  
PR 16-FEB-2001; 2001US-0269535P.  
XX  
PA (BRIM ) BRISTOL-MYERS SQUIBB CO.  
XX  
PI Feder J, Lee L, Chen J, Jackson DG, Ramanathan C, Siemers N;  
PI Chang H;  
XX  
DR WPI; 2002-674925/72.  
XX  
DR P-PSDB; AA019187.  
XX  
PT New isolated nucleic acid molecules encoding human glycine receptor A4  
PT (HGRA4) polypeptides, useful for preventing, treating and ameliorating  
PT conditions, e.g. neural or gastrointestinal disorders.  
XX  
PS Claim 1; Fig 2; 349pp; English.  
XX  
CC The present invention provides the protein and coding sequences of the  
human glycine receptor alpha 4 (HGRA4) and its splice variant HGRA4sv.



PD		29-AUG-2002.
XX		
PF		13-FEB-2002; 2002WO-US004329.
XX		
PR		16-FEB-2001; 2001US-0269535P.
XX		
PA	(BRIM ) BRISTOL-MYERS SQUIBB CO.	
XX		
Feder J,	Lee L, Chen J, Jackson DG, Ramanathan C, Siemens N;	
PI	Chang H;	
XX		
DR	WPI; 2002-674925/72.	
DR	P-PSDB; AAO19186.	
XX		
PT	New isolated nucleic acid molecules encoding human glycine receptor A4	
PT	(HGRA4) polypeptides, useful for preventing, treating and ameliorating	
PT	conditions, e.g. neural or gastrointestinal disorders.	
XX		
XX	Claim 14; Fig 1; 349pp; English.	
PS		
CC	The present invention provides the protein and coding sequences of the	
CC	human glycine receptor alpha 4 (HGRA4) and its splice variant HGRA4sv.	
CC	The sequences can be used in the treatment of neural disorders,	
CC	gastrointestinal disorders, disorders related to hyperglycine receptor	
CC	activity, cardiovascular disorders, reproductive disorders, or bacterial,	
CC	viral and parasitic infections. The present sequence is a coding sequence	
CC	of the invention	
XX		
SQ	Sequence 2565 BP; 596 A; 681 C; 590 G; 698 T; 0 U; 0 Other;	
	Query Match 93.0%; Score 1524.8; DB 6; Length 2565;	
	Best Local Similarity 97.3%; Pred. No. 0;	
	Matches 1578; Conservative 0; Mismatches 2; Indels 42; Gaps 1	
Qy	1 ATGACAACCTCTGTTCCTGGCAACCTCTCCTTCTCTCTCTCGACCCTGCCAGGGCAG 60	
Dd	1 ATGACAACCTCTGTTCCTGGCAACCTCTCCTTCTCTCTCTCGACCCTGCCAGGGCAG 60	
Qy	61 GTCTCTCCTCAGGGTGCCCTTGGCAAAAGAGGAAGTCAAATCTGGAAACCAAGGGGTCCCAG 120	
Dd	61 GTCTCTCCTCAGGGTGCCCTTGGCAAAAGAGGAAGTCAAATCTGGAAACCAAGGGGTCCCAG 120	
Qy	121 CCCATGTCCCCCTCTGATTTCTTAGACAAAATTATGGGGCGAACACTCTGGATATGATGCC 180	
Dd	121 CCCATGTCCCCCTCTGATTTCTTAGACAAAATTATGGGGCGAACACTCTGGATATGATGCC 180	
Qy	181 AGGATTCGGCCCAATTTTAAAGGCCCCACCGTGAACGTGACCTGCAACATCTTCATCAAC 240	
Dd	181 AGGATTCGGCCCAATTTTAAAGGCCCCACCGTGAACGTGACCTGCAACATCTTCATCAAC 240	
Qy	241 AGTTTCAGCTCCCTCACCAAGACCAATGGACTACCGGGTGAATGTCTTTCTGGGGCAA 300	
Dd	241 AGTTTCAGCTCCGTACCAGACCAATGGACTACCGGGTGAATGTCTTTCTGGGGCAA 300	
Qy	301 CAGTGGAAATGACCCACGCCTGTCTTACCAGAAATACTCTGATGACTCTCTGGACCTCGAT 360	
Dd	301 CAGTGGAAATGACCCACGCCTGTCTTACCAGAAATATCTGATGACTCTCTGGACCTCGAT 360	
Qy	361 CCCTCCATGTGGACTTATCTGGAAGCCAGACCTCTTCTTTGCTTAATGAGAAGGGGCC 420	
Dd	361 CCCTCCATGTGGACTTATCTGGAAGCCAGACCTCTTCTTTGCTTAATGAGAAGGGGCC 420	
Qy	421 AAC TTCATGAGGTGACCAACGGAACAAAGTTACTTGC GCATCTTCCAAGAA TGGGAATGTG 480	
Dd	421 AAC TTCATGAGGTGACCAACGGAACAAAGTTACTTGC GCATCTTCCAAGAA TGGGAATGTG 480	
Qy	481 CTGTACAGCATCAGGCTGACCCCTCATTTTGTGCTGCTGCTGATGGACCTCAAGAAC TTCCCC 540	
Dd	481 CTGTACAGCATCAGGCTGACCCCTCATTTTGTGCTGCTGCTGATGGACCTCAAGAAC TTCCCC 540	
Qy	541 ATGGACATCCAGAGCTGACAGTGACAGTGACGTGTGAGAGCTCATTCAT ACTCTG CAGCCCTCTG 600	
Dd	541 ATGGACATCCAGAGCTGACAGTGACAGTGACGTGTGAGAGCTCATTCAT ACTCTG CAGCCCTCTG ----- 577	

Qy	601	CCATCTCTGTCACTTTTCAGTGTGGCTACACATGAAGAACCCTCGTGTGTAGTGGCTGGAA	660
Db	578	-----TTGGCTACACCATGAAGAACCCTCGTGTGTAGTGGCTGGAA	618
Qy	661	GATGCTCTCTGCTGTCCAAGTGGCTGAGGGGCTGACTCTGCCCCAGTTTATCTTTGGCGGAT	720
Db	619	GATGCTCTCTGCTGTCCAAGTGGCTGAGGGGCTGACTCTGCCCCAGTTTATCTTTGGCGGAT	678
Qy	721	GAGAAGGATCTTAGCTGTGTGTACCAAGCACTACAAACACAGGGAAATTCACCTGCATCGAG	780
Db	679	GAGNAGGATCTTAGCTGTGTGTACCAAGCACTACAAACACAGGGAAATTCACCTGCATCGAG	738
Qy	781	GTAAAGTTTCACTGGAAACGGCAGATGGGCTACTATCTGATTCAGATGTATCATCCCCAGC	840
Db	739	GTAAAGTTTCACTGGAAACGGCAGATGGGCTACTATCTGATTCAGATGTATCATCCCCAGC	798
Qy	841	CTACTCATCTCATCTGCTCTGGGTCTCTTCTGGATCAACATGGATGCTGCCCTGCC	900
Db	799	CTACTCATCTCATCTGCTCTGGGTCTCTTCTGGATCAACATGGATGCTGCCCTGCC	858
Qy	901	CGTGTGGGCTGTGGSCATCACACCGTGCTCACCATGACCACCCAGAGCTCTGGCTCCCGG	960
Db	859	CGTGTGGGCTGTGGSCATCACACCGTGCTCACCATGACCACCCAGAGCTCTGGCTCCCGG	918
Qy	961	GCCTCTTTTGCTTAAGGTGTCTACGTGAAGGCAATCGACATCTGGATGGCTGTGTGTCTG	1020
Db	919	GCCTCTTTTGCTTAAGGTGTCTACGTGAAGGCAATCGACATCTGGATGGCTGTGTGTCTG	978
Qy	1021	CTCTTTGTGTTCGTGTGCTGTGGAGTATGCTGCCATAAATTTTGTTCCTGTCAGCAT	1080
Db	979	CTCTTTGTGTTCGTGTGCTGTGGAGTATGCTGCCATAAATTTTGTTCCTGTCAGCAT	1038
Qy	1081	AAAGAAATTCATACGACTTCCGAAGACGACAGCGCCACCGCTTGGAGGAAGATATCATC	1140
Db	1039	AAAGAAATTCATACGACTTCCGAAGACGACAGCGCCACCGCTTGGAGGAAGATATCATC	1098
Qy	1141	CAAGAAAGTCGTTTTCTATTTCCGTGGCTATGGCTTGGGCCACCTGCCTGCAGGCAAGAGAT	1200
Db	1099	CAAGAAAGTCGTTTTCTATTTCCGTGGCTATGGCTTGGGCCACCTGCCTGCAGGCAAGAGAT	1158
Qy	1201	GGAGGTCCAATGGAAGTTCTGGCATTTATAGTCCCAACCTCAGGCCCTCTTCTTAAGG	1260
Db	1159	GGAGGTCCAATGGAAGTTCTGGCATTTATAGTCCCAACCTCAGGCCCTCTTCTTAAGG	1218
Qy	1261	GAAGGAAACCAACGCGGAACCTCTACGTGGACTGAGCCAAAGAAATTGACACCATCTCC	1320
Db	1219	GAAGGAAACCAACGCGGAACCTCTACGTGGACTGAGCCAAAGAAATTGACACCATCTCC	1278
Qy	1321	CGGCTGTCTTCCCTTTCACTTTCTCATCTTCAATATCTTTACTGGGTGTCTATAAA	1380
Db	1279	CGGCTGTCTTCCCTTTCACTTTCTCATCTTCAATATCTTTACTGGGTGTCTATAAA	1338
Qy	1381	GTGCTATGGTCAGAAAGATATCCACAGGCTCTGTGAATAGGGTGGAGCTATAGAGTCCT	1440
Db	1339	GTGCTACGGTCAGAAAGATATCCACAGGCTCTGTGAATAGGGTGGAGCTATAGAGTCCT	1398
Qy	1441	GCTGCTGGCCTCTGCTTCTCTCTGGGTGGGCTTTCTCCCTCAGTTAGACTCCATTAGGG	1500
Db	1399	GCTGCTGGCCTCTGCTTCTCTCTGGGTGGGCTTTCTCCCTCAGTTAGACTCCATTAGGG	1458
Qy	1501	GTTTTGGACAGTTTCTCTCTGATCTCCCACTCAGAACTTCAACTACCAAGTCCCAAGCTAT	1560
Db	1459	GTTTTGGACAGTTTCTCTCTGATCTCCCACTCAGAACTTCAACTACCAAGTCCCAAGCTAT	1518
Qy	1561	GTGGGCTTATATGTCATGGTGCCAATGGTGGCTGTACTTATAAGATGGCTTATCTACCC	1620
Db	1519	GTGGGCTTATATGTCATGGTGCCAATGGTGGCTGTACTTATAAGATGGCTTATCTACCC	1578
Qy	1621	TA 1622	
Db	1579	TA 1580	

1621 TA 1622

1579 TA 1580











DT	20-MAY-2004	(first entry)
XX	DE	Human TRICH-38 coding sequence, SEQ ID 86.
KW	XX	Human; Transporters and Ion Channels protein; TRICH; Transporter;
KW	XX	Ton Channel; transport disorder; neurological disorder; muscle disorder;
KW	XX	Immunological disorder; cell proliferative disorder; infection;
KW	XX	Neuroprotective; Muscular; Respiratory; Antidiabetic; Antiparkinsonian;
KW	XX	Nootropic; Anticonvulsant; Anthithyroid; Ophthalmological; Antibacterial;
KW	XX	Antiinflammatory; Cardiant; Antiangiinal; Antiarrhythmic; Antiasthmatic;
KW	XX	Endocrine; Hypotensive; Anti-HIV; Antianaemic; Antiarteriosclerotic;
KW	XX	Dermatological; Antiarthritic; Osteopathic; Antipsoriatic; Antirheumatic;
KW	XX	Immunosuppressive; Anticancer; Gastrointestinal; Hepatotrophic; Cytostatic;
KW	XX	Virucide; Fungicide; Antiparasitic; Protozoacide; Antihelminthic;
KW	XX	Gene Therapy; TRICH-38; gene; ds.
OS	XX	Homo sapiens.
XX	XX	WO2004013293-A2.
PX	PN	12-FEB-2004.
PD	XX	01-AUG-2003; 2003WO-US024164.
PF	XX	02-AUG-2002; 2002US-0400953P.
PR	XX	12-AUG-2002; 2002US-0403321P.
PR	PR	19-AUG-2002; 2002US-0404693P.
PR	PR	29-AUG-2002; 2002US-0407062P.
XX	PA	(INCY-) INCYTE CORP.
XX	PI	Richardson TW, Hafalia AJA, Khare R, Tran UK, Emerling BM;
PI	PI	Lee SY, Chawla NK, Elliott VS, Swarnakar A, Chang H, Kable AB;
PI	PI	Becha SD, Mason PM, Jiang X, Jackson AA, Marquis JP, Das D;
PI	PI	Gietzen KJ, Bing L, Baughn MR, Lee EA, Thornton MB, Borowsky ML;
PI	PI	Harland L, Arvizu CS, Yao MG, Gandhi AR, Au-Young JK, Warren BA;
PI	PI	Lal PG, Lindquist EA;
DR	XX	-WPI; 2004-180430/17.
DR	XX	P-PSDB; ADJ27206.
DR	XX	New human transporters and ion channels and nucleic acids encoding them,
PT	PT	useful for diagnosing, treating or preventing cystic fibrosis, Becker's
PT	PT	muscular dystrophy, Parkinson's disease, myasthenia gravis, prion disease
PT	PT	or cancers.
PS	XX	Claim 5; SEQ ID NO 86; 357pp; English.
XX	CC	The present invention relates to novel human Transporters and Ion
CC	CC	Channels proteins referred to collectively as TRICH (TRICH-1 - TRICH-48,
CC	CC	ADJ27169-ADJ27216) and their coding sequences (ADJ27217-ADJ27264). The
CC	CC	TRICH proteins and coding sequences are useful in the diagnosis,
CC	CC	treatment and prevention of transport, neurological, muscle,
CC	CC	immunological or cell proliferative disorders, e.g. akinesia, amyotrophic
CC	CC	lateral sclerosis, ataxia telangiectasia, cystic fibrosis, Becker's
CC	CC	muscular dystrophy, Bell's palsy, diabetes mellitus, diabetes insipidus,
CC	CC	diabetic neuropathy, Parkinson's disease, malignant hyperplasia,
CC	CC	myasthenia gravis, peripheral neuropathy, Alzheimer's disease,
CC	CC	Huntington's disease, Grave's disease, cataracts, bacterial and viral
CC	CC	meningitis, brain abscess, prion disease, Creutzfeldt-Jakob disease,
CC	CC	fatal familial insomnia, neurofibromatosis, Down syndrome,
CC	CC	cardiomyopathy, myocarditis, myotonic dystrophy, lipid myopathy, angina,
CC	CC	arhythmias, asthma, Cushing's syndrome, hypertension, hypoglycaemia,
CC	CC	epilepsy, lactic acidosis, AIDS, anaemia, adult respiratory distress
CC	CC	syndrome, atherosclerosis, bronchitis, Crohn's disease, atopic
CC	CC	dermatitis, erythroblastosis fetalis, osteoarthritis, osteoporosis,
CC	CC	pancreatitis, psoriasis, rheumatoid arthritis, scleroderma, systemic
CC	CC	lupus erythematosus, ulcerative colitis, cirrhosis, cancers, viral,
CC	CC	bacterial, fungal, parasitic, protozoal or helminthic infections.
XX	XX	Sequence 1224 BP; 284 A; 349 C; 297 G; 294 T; 0 U; 0 Other;
SQ	XX	Query Match            57.8%; Score 947.8; DB 12; Length 1224;





Qy	481	CTGTACAGCATCAGGCTGACCCCTCATTTTGTCTGCTGTAGTGGACCTCAAGAACTTCGCC	540
Db	773	CCGTACAGCATCAGGCTGACCCCTCATTTTGTCTGCTGTAGTGGACCTCAAGAACTTCGCC	832
Qy	541	ATGAGACATCCAGAGCTGCACGATCGAGCTTGAGAGGCTCATCCATACCTGCGAGCCCTCTG	600
Db	833	ATGAGACATCCAGAGCTGCACGATCGAGCTTGAGAGCT-----	869
Qy	601	CCATCTCTGTCACTTTTCACTTGGCTTACCAATGAAGAACCTCTGTTTGTAGTGGCTGGAA	660
Db	870	-----TTGGCTACACCATGAAGAACCTCTGTTTGTAGTGGCTGGAA	910
Qy	661	GATGCTCCTGCTGTCGAAGTGGCTGAGGGCTGACTCTGCCCCAGTTTATCTTGGCGGAT	720
Db	911	GATGCTCCTGCTGTCGAAGTGGCTGAGGGCTGACTCTGCCCCAGTTTATCTTGGCGGAT	970
Qy	721	GAGAAGGATCTAGGCTGTGTGTACCAAGCACTTACAACACAGGGAAATTCACCTGTATCGAG	780
Db	971	GAGAAGGATCTAGGCTGTGTGTACCAAGCACTTACAACACAGGGAAATTCACCTGTATCGAG	1030
Qy	781	GTAAAGTTTACCTGGAAACGGCAGATGGGCTACTATCTGTATTCAGATGTACATCCCGAGC	840
Db	1031	GTAAAGTTTACCTGGAAACGGCAGATGGGCTACTATCTGTATTCAGATGTACATCCCGAGC	1090
Qy	841	CTACTCATCGTCATCCTGTCTCTGGGTCTCCTTCTGGGATCAACATGGATGCTGCCCTGCCC	900
Db	1091	CTACTCATCGTCATCCTGTCTCTGGGTCTCCTTCTGGATCAACATGGATGCTGCCCCTGCCC	1149
Qy	901	CGTGTGGGCTTGGGCATCAACACCGTG	927
Db	1150	CGTGTGGGCTTGGGCATCAACACCGTG	1176

## RESULT 8

AAAL49665	
ID	AAAL49665 standard; DNA; 993 BP.
XX	
XX	AAAL49665;
XX	
XX	27-NOV-2002 (first entry)
XX	
DE	Human glycine receptor alpha subunit HGRA4 related polynucleotide.
XX	
XX	Human; glycine receptor alpha subunit 4; HGRA4; HGRA4sv; splice variant;
KW	cardiovascular disorder; reproductive disorder; neural disorder;
KW	cardiant; antiarrhythmic; antianginal; antiidiarrheic; antiulcer;
KW	nootropic; neuroprotective; antibacterial; virucide; protozoacide;
KW	nervous system disorder; gastrointestinal disorder; gene therapy;
KW	infection; ds.
XX	
XX	Homo sapiens.
OS	
XX	
XX	WO200266606-A2.
PN	
XX	
XX	29-AUG-2002.
PD	
XX	
XX	13-FEB-2002; 2002WO-US004329.
PF	
XX	
XX	16-FEB-2001; 2001US-0269535P.
PR	
XX	
XX	(BRIM ) BRISTOL-MYERS SQUIBB CO.
PA	
PA	
XX	
PI	Feder J, Lee L, Chen J, Jackson DG, Ramanathan C, Siemens N;
PI	Chang H;
XX	
XX	WPI; 2002-674925/72.
DR	
XX	
XX	New isolated nucleic acid molecules encoding human glycine receptor A4
PT	(HGRA4) polypeptides, useful for preventing, treating and ameliorating
PT	conditions, e.g. neural or gastrointestinal disorders.
PT	
XX	
XX	Disclosure; Fig 7; 349pp; English.
PS	
XX	



Db 1042 GCTACTACCTGATTCAGATGATATATTCCTCCAGCGCTGCTCATTTGTCATCTCTCATGGATCT 1101  
QY 869 CTTCTTGGATCAACATGGATGCTGCTCCCTGCGCGTGGGCTGGGCATCACCACCGTGC 928  
Db 1102 CTTCTTGGATCAACATGGATGCTGCTCCCTGCGCGTGGGCTGGGCATCACCACCGTGC 1161  
QY 929 TCACCATGACCCACAGAGCTGCTGCTCCCGGCTCTTTTGGCTTAAGGTGTCTACGTGA 988  
Db 1162 TCACCATGACCCACAGAGCTGCTGCTCCCGGCTCTTTTGGCTTAAGGTGTCTACGTGA 1221  
QY 989 AGGCATGACATCTGGATGGCTGTGTGCTCTTTTGGCTTAAGGTGTCTACGTGA 1048  
Db 1222 AAGCTATTGACATCTGGATGGCTGTGTGCTCTTTTGGCTTAAGGTGTCTACGTGA 1281  
QY 1049 ATGCTGCCATAAATTTTGTTCGTCAGCATAAAGATTTTATACACTTCCGAAGAAGC 1108  
Db 1282 ATGCCGCTGCACTTTTGTCTCGCAACATTAAGAGCTGCTCCGATTCAGAGGAAGC 1341  
QY 1109 AGAGGCGCAACGCTTGGAGGAGATATCATCAAGAAAGTCTTTTCTATTTTCCGTGGCT 1168  
Db 1342 GGAGACATCACAAAGGAGGATGAAGCT-----GGAGAGGCGCTTTAACTTCTGTGCT 1395  
QY 1169 ATGGCTTGGCC---ACTGCTGCGAGCAAGAGATGAGGTCCAATGGAGGTC---TG 1222  
Db 1396 ATGGGATGGGCGGAGCTGTCTACAGGCCAAGAGTGGCATCTCAGTCAAGGGCGCCAACA 1455  
QY 1223 GCATTTATAGTCCCAACTCCAGCCCTCTTCTAAGGGAAGAGAAACACAGCGGAAC 1282  
Db 1456 ACATGAACACCAACCCCTCTCTGACCATCTAAGTCCCAGAGGAGATCGGAAC 1515  
QY 1283 TCTACGTGACTGAGCCAGAGAAATTTGACACCATCTCCCGGGCTGCTTCCCTTTCACAT 1342  
Db 1516 TCTTATCAGAGGCGCAAGAGATCGACAAATATCCCGATTTCCCGATGGCT 1575  
QY 1343 TCTCATCTCATATCTTCTACTGCTGTTCTATTAAGTGTATGTCAGAGATATCC 1402  
Db 1576 TCTCATTTTCAACATGTTCTACTGGATCATCTACAAGATTGTCGTAGAGAGGACGTCC 1635  
QY 1403 ACCA 1406  
Db 1636 ACNA 1639

RESULT 10  
ID ADQ89057  
XX ADQ89057 standard; cDNA; 1715 BP.  
XX AC ADQ89057;  
XX DT 21-OCT-2004 (first entry)  
XX DE Human urological disorder related protein 884 encoding cDNA SEQ:9.  
XX KW urological disorder; uropathic; cytostatic; urinary incontinence;  
XX KW benign prostatic hyperplasia; human; Gene; ss.  
XX OS Homo sapiens.  
XX PH Key Location/Qualifiers  
XX FT CDS 297..1646  
XX FT /\*tag= a  
XX FT /product= "urological disorder related protein 884"  
XX WO2004065576-A2.  
XX PN 05-AUG-2004.  
XX PD 14-JAN-2004; 2004WO-US000750.  
XX PF 15-JAN-2003; 2003US-0440318P.  
XX PR 04-FEB-2003; 2003US-0444783P.  
XX PR 27-MAR-2003; 2003US-0457901P.  
XX PR 08-MAY-2003; 2003US-0468775P.

PR 19-MAY-2003; 2003US-0471614P.  
PR 16-JUN-2003; 2003US-0478742P.  
PR 18-JUL-2003; 2003US-0488529P.  
PR 30-JUL-2003; 2003US-0491156P.  
PR 02-SEP-2003; 2003US-0499594P.  
PR 26-SEP-2003; 2003US-0508332P.  
XX (MILL-) MILLENNIUM PHARM INC.  
XX PA Karicheti V, Silos-Santiago I, Eliasof SD;  
XX PI WPI; 2004-562167/54.  
XX DR P-PSDB; ADQ89058.  
XX PT Use of polypeptides related to urological disorders, e.g. 44390, 54181,  
PT 211 or for identifying a compound capable of treating a urological  
PT disorder or identifying and treating a subject having a urological  
PT disorder.  
XX Claim 1; SEQ ID NO 9; 542pp; English.  
XX CC The present invention describes the use of polypeptides related to  
CC urological disorders for identifying a compound capable of treating a  
CC urological disorder, identifying a subject having a urological disorder,  
CC or treating a subject having a urological disorder. Also described: (1) a  
CC method for identifying a compound capable of treating a urological  
CC disorder; (2) a method for identifying a subject having a urological  
CC disorder; and (3) a method for treating a subject having a urological  
CC disorder. The compound has uropathic and cytostatic activities. The  
CC polypeptides related to urological disorders are useful for identifying a  
CC compound capable of treating a urological disorder, identifying a subject  
CC having a urological disorder, or treating a subject having a urological  
CC disorder. Disorders include urinary incontinence and benign prostatic  
CC hyperplasia. The present sequence encodes a human urological disorder  
CC related protein, which is used in the exemplification of the present  
CC invention.  
XX Sequence 1715 BP; 424 A; 480 C; 407 G; 404 T; 0 U; 0 Other;

Query Match 40.0%; Score 656; DB 13; Length 1715;  
Best Local Similarity 71.9%; Pred. No. 4.5e-177; Indels 57; Gaps 5;  
Matches 952; Conservative 0; Mismatches 315;  
QY 89 AGGAAGTCAAATCTGGAAACCAAGGGGTCCAGCCATGTCCCGCTCTGATTTCTTAGACA 148  
Db 367 AGGAGCTGAAGTGTCTGCTCCGCAACCAAGCTATGTACCTCGGATTTCTTGATA 426  
QY 149 AACCTATATGGGCGAATCTGGATATGATGCGAGGATTCGGCCCAATTTTAAAGGCCAC 208  
Db 427 AGCTAATGGGGAACCTCCGGATATGATGCCAGGATCAGGCCCAATTTTAAAGGTCCCC 486  
QY 209 CCGTGAACGTGACCTGCAACATCTTCATCAACAGTTTTCAGTCCGTCACCAAGACCACA 268  
Db 487 CAGTGAACGTGAGCTGCAACATTTTCAACAGCTTTGGTCCATATGCTGAGAACACA 546  
QY 269 TGGACTACCGGCTGAATGTCTTCTTCGCGCAACAGTGAATGACCCACGCTGTCTTACC 328  
Db 547 TGGACTATAGGCTCAACATCTTCTTCGCGCAGCAATGGAACGCCCGCTGCTATA 606  
QY 329 GAGATATCTGATGACTCTCTGGACCTCGATCCCTCCATGCTGAGTCTCTATCTGAAGC 388  
Db 607 ATGAATACCTGACGACTCTCTGGACCTGGACCCATCCATGCTGGAATCCATCTGGAAC 666  
QY 389 CAGACCTCTTCTTGTCTAATGAGAGGGGCGCAACTTCCATGAGGTGACCGACAACA 448  
Db 667 CTGACCTGTCTTGTGCAACAGAGAGGGGCGCAACTTCCATGAGATCACCACACAACA 726  
QY 449 AGTTACTGCGCATCTTCAAGAATGGGAATGTGTGTACAGCATCAGGCTGACCTCATTT 508  
Db 727 AATTGCTAAGGATCTCCCGGAATGGGAATGTCTCTACAGCATCAGATCACCCTGACAC 786  
QY 509 TGTCTGCTGATGAGACTCTCAAGAACTTCCCATGGAACATCCAGAGCTGACGATCAGC 568



QY 444 CAACAAAGTACTGCGCATCTTCAAGAAATGGAAATGTGCTGTACAGCATCAGGCTGACCTT 503  
 DB 1616 TACAAAGTCTGCGGATTTCCAAATATGGCAAGTGCTCTACAGTATTAGACTCACCTT 1675  
 QY 504 CATTTTGTCTGCTGATGGAAGCTCAAGAACTTCCCATGGACATCCAGACGTGACAGAT 563  
 DB 1676 GACTTTATCTGTGCCATGGACCTGAAGAACTTCCAAATGGATGTCCAGACCTGTACAAT 1735  
 QY 564 GCAGCTTGGAGCTCATCATCTCTGAGCCCTCTGCCATCTCTGTCTACTTTCAGTTGG 623  
 DB 1736 GCAGCTGGAGATT-----TTGG 1753  
 QY 624 CTACACCATGAAGACCTCGTGTGTTGAGTGGCTGGAAGATGCTCTCTGTCTCAAGTGC 683  
 DB 1754 GTACACCATGAATGACCTGATATTGAGTGGTTAAGTATGGTC---CAGTACAGTTGC 1810  
 QY 684 TGAGGGCTGACTTGTGCCCCAGTTTATCTTGGGGATGAGAGGATCTAGGCTGTGTGAC 743  
 DB 1811 TGAAGGACTCACCTGCTCAGTTTATTTTGAAGAGAGAGAACTTGGCTATTGGCAC 1870  
 QY 744 CAAGCACTACACAGAGGAATTCACCTGCATCGAGTTAAAGTTTCACCTGGAACGCA 803  
 DB 1871 AAAGCATTACACACTGGAAGTTTACCTGCAATTGAGGTCAAGTTTCACCTGGAACGCA 1930  
 QY 804 GATGGCTACTTATCTGATTTCAGATGTACATCCCAGCCTACTCATCGTCTATCTGCTCG 863  
 DB 1931 GATGGCTATTATTGATTCAGATGTATATCCCAGCCTGCTGATAGTCAATTTGTCTG 1990  
 QY 864 GGTCTCTTCTGATCAACATGATGCTGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 923  
 DB 1991 GGTCTCTTCTGATCAACATGATGCTGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2050  
 QY 924 CGTGTCTACCATGACCAACAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 983  
 DB 2051 CGTCTCTGCAATGACTACCCAGAGTTTCAGGTTTCCAGGGCATCTCTGCCAAAGGCTCTCAT 2110  
 QY 984 CGTGAAGGCAATCGACATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1043  
 DB 2111 TGTGAAGCAATGACATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2170  
 QY 1044 GGAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1103  
 DB 2171 GGAATATGAGCAAGTGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2230  
 QY 1104 AAGGCAAGGCGCAAGCTTGGAGGAGATATCATCCAGAAAGTCTGTTCTATTTCG 1163  
 DB 2231 ACGACAGAGAGGCGAATAAGGAAGAGATGTTACTCTGGAAGCGGTTTTAACTTCAG 2290  
 QY 1164 TGGCTATGCTTGGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1223  
 DB 2291 CGGTTATGGATGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2347  
 QY 1224 CATTTATAGTCCCAACCTTCCAGCCCTCTTCTTAAGGGAAGGAGAAACCAACGCGGAACT 1283  
 DB 2348 -----ACCTGCCAACCTTCCGCAACCCCAAGAGATGCAAGTCTATCAAGAAGAA 2401  
 QY 1284 CTAAGTGGATGAGCAAGAGATTCACACCATCTCCGGGCTGCTTCTCCCTTTCACTTT 1343  
 DB 2402 GTTTGTGGATTCGGGCAAAAGAAATGACACCATATCTCGAGCTGCTTCCCACTGGCCCTT 2461  
 QY 1344 CCTCATCTTCAATATCTTCTACTGGGTGTGCTATAAGTGTCTATGCTGAGAAGATATCA 1403  
 DB 2462 CCTCATTTCAACATCTTTTACTGGATCACAATCAAGATCATTCGGCATGAAGATGTCCA 2521  
 QY 1404 CCAG 1407  
 DB 2522 CAAG 2525

RESULT 12  
 AB218534  
 ,ID AB218534 standard; cDNA; 1857 BP.

XX AB218534;  
 AC 23-JAN-2003 (first entry)  
 DE Group III cDNA cancer related clone SEQ ID NO:960.  
 KW Human; cancer; tumour; therapy; diagnosis; CT antigen; CP antigen;  
 KW immune response; virology; immunology; microbiology; molecular biology;  
 KW recombinant DNA technology; gene; ss.  
 XX Homo sapiens.  
 OS WO200278516-A2.  
 XX 10-OCT-2002.  
 XX 28-MAR-2002; 2002WO-US010421.  
 XX 30-MAR-2001; 2001US-0280255P.  
 PR 28-AUG-2001; 2001US-0315563P.  
 PR 09-JAN-2002; 2002US-034713P.  
 XX (CORI-) CORIXA CORP.  
 XX Wang T, Wang S, Bangur CS, Gaiger A;  
 PI WPI; 2003-058387/05.  
 XX New immunogenic polynucleotides or polypeptides useful for diagnosing,  
 PT preventing and treating cancer expressing CT or CP mRNA antigens, and in  
 PT virology, immunology, microbiology, molecular biology and recombinant DNA  
 PT techniques.  
 XX Claim 1; SEQ ID NO 960; 207pp; English.  
 PS ABQ17575 to ABQ20506 represent isolated polynucleotide (I) sequences, and  
 CC ABP54446 to ABP54472 represent protein (II) sequences, from the present  
 CC invention. (I) and (II) have cytostatic activity and can be used in gene  
 CC therapy and vaccines. (I), (II), antibodies and compositions from the  
 CC present invention are useful for diagnosing, preventing and treating  
 CC cancer, which expresses CT or CP mRNA antigens. They are useful for  
 CC stimulating immune response. They can also be useful in virology,  
 CC immunology, microbiology, molecular biology and recombinant DNA  
 CC techniques. N.B. The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX Sequence 1857 BP; 526 A; 416 C; 417 G; 498 T; 0 U; 0 Other;

Query Match 39.6%; Score 650.2; DB 8; Length 1857;  
 Best Local Similarity 70.1%; Pred. No. 2.2e-175;  
 Matches 943; Conservative 0; Mismatches 348; Indels 54; Gaps 3;  
 QY 63 CCTCTCTCAGGGTGGCTTGGCAAAAGAGGAAGTCAAATCTGGAACCAAGGGGTCCAGCC 122  
 DB 450 CTTTCAGGCGCTTCTGCAAGACCATGACTCCAGGCTCTGGAACCAACCTTCACAGAC 509  
 QY 123 CATGTCCCCTCTGATTTCTTAGACAACTTATGGGCGAATCTGGATATGATGCCAG 182  
 DB 510 CCTATCTCTCTCAGATTTCTTGGACAAAGTAAATGGGAAGGACATCAGGATATGATGCAAG 569  
 QY 183 GATTGGCCCAATTTTAAAGGCCCCCTGGAACGTGACCTGCAACATCTTCAATCAACAG 242  
 DB 570 AATCAGGCCCAATTTTAAAGGTCCTCCAGTAAAGTCTTTCGATATTTTATCAACAG 629  
 QY 243 TTTTCAGCTCCGTCAACAGACCAATGACCTACCGGGTGAATGCTTCTTGGCGCAACA 302  
 DB 630 TTTTGGATCAGTCACAGAAACGACATGACTACCGAGTGAATATTTTCTGAGACAACA 689  
 QY 303 GTGGAATGACCAACGCTCTCTACCGAATATCTGATGACTCTCTGGACCTCATCC 362  
 DB 690 GTGGAATGATTCACGGCTGGCGTACAGTGAGTACCCAGATGACTCTCCCTGGACTTGGACCC 749



QY 363 CTCCATCTGGACTCTATCTGGAGCCAGACCTCTTCTTTGCTAATGAGAAAGGGCCAA 422  
 Db 750 ATCCATGCTAGACTCCTATTTGGAAACCAAGATTTGTTCTTTGGCAATGAGAGGGTGCCAA 809  
 QY 423 CTTCATAGGCTGACCAAGCAACCAAGTTACTGCGCATCTTCAAGAAATGGGAATGTCT 482  
 Db 810 CTTCACAGATGACCACTGACACAAATGCTACGGATTTGAAATATGGCAAGTGT 869  
 QY 483 GTACAGCATGAGCTGACCTCATTTTGTCTGCTGATGGACCTCAAGAACTTCCCAT 542  
 Db 870 CTACAGTATCAGACTCACCTTGACCTTATCTGCTCCATGGACTTGAAGAACTTTCCGAT 929  
 QY 543 GGACATCCAGAGCTGACCATGCTGAGCTTGAAGCTCATCTACTCTGAGCCCTCTGCC 602  
 Db 930 GGATGCTCCAGACTGTACAAATGCTGAGAGTT----- 964  
 QY 603 ATCTGTCTACCTTTAGTTGGCTTACACCATGAAGACCTCGTGTGTTGAGTGGCTGGAAGA 662  
 Db 965 -----TTGGGTACACGATGATGACCTGATATTTGAGTGGTTAAGTGA 1007  
 QY 663 TGTCTCTGTCTCAAGTGGCTGAGGGCTGACTCTGCCCCAGTTTATCTTGGGGATGA 722  
 Db 1008 TGGTC---CAGTCAAGTTGCTGAAGCAATTGACCTGCCCCAGTTTATTTTGAAGAAGA 1064  
 QY 723 GAAGATCTAGCTGTTGTACACAGCACTACACACAGGGAATTCACCTGCAATCGAGT 782  
 Db 1065 GAAGGAATTTGGCTACTGTACAAAGCACTACACACTGGAAGATTTTACCTGCAATGAGT 1124  
 QY 783 AAGTTTACCTGGAAGCGCAGATGGCTACTATCTGATTCAGATGATACATCCCAAGCT 842  
 Db 1125 CAAAGTTTACCTGGAAGCGCAATGGGATATTTTATTCAGATGATACATCCCAAGCT 1184  
 QY 843 ACTCATGCTCATCTCTCTGCTGCTCTCTTCTGATCAACATGATGCTGCCCTCCCG 902  
 Db 1185 GCTATAGTAATTTGCTGCTGCTCTCTTGGATAAATGATGCGAGCCCTGCCAG 1244  
 QY 903 TGTGGCCCTGGGATCACCACCGTGTCTACATGACCCAGAGCTTGGCTCCCGGGC 962  
 Db 1245 GGTGCACTGGGATCACCACAGCTTAAAGATGACCCAGAGTTTCAAGGCTCCAGGGC 1304  
 QY 963 CTCCTTGGCTAGGCTCTCTAGTGAAGCAATCGACATCTGATGCTGCTGCTGCT 1022  
 Db 1305 ATCTGCAAGGCTCTCTATGTAAGGATGATGATCTGATGCGGCTGCTCTCT 1364  
 QY 1023 CTTTGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1082  
 Db 1365 GTTGTGTTGCTGCTTACTGGAATACGAGCGGTGACTTCTCTCCAGCAACACAA 1424  
 QY 1083 AGAATTCATAGCTTGAAGAGGAGGAGGCGCCAAACGCTTGGAGGAAGATATATCA 1142  
 Db 1425 GGAGTTCTGCGCTCCGAAGAAGACAGAGAGGCGAATAAGGAAGAAGAGCTTACTCG 1484  
 QY 1143 AGAAGTCTGTTCTATTTCCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1202  
 Db 1485 TGAAGTCTGTTTAAATTTTACGCGTTTATGGATGGGTCTACTGCTCAAAGTGAAGATG 1544  
 QY 1203 AGTCCAAATGGAAGGTTCTGGCATTTATAGTCTCCCAACCTCCAGCCCTCTTCTAAGGA 1262  
 Db 1545 AACAGTGTCAAGGCCAC-----ACCTGCCAACCCACTCCCAACACCGCCAAAGA 1595  
 QY 1263 AGAGAAACACCGGAAACTCTACTGATGAGCAAGAGAAATTTGACCACTTCCCG 1322  
 Db 1596 TGGAGATGCTATCAAGAAGAGTTTGTGGACCGGGCAAAAGATTGACACGATATCTCG 1655  
 QY 1323 GGTGTGTTCCCTTTCACCTTCTCTATCTTCAATATCTTACTGCTGCTGCTGCTGCT 1382  
 Db 1656 AGTGTCTTCCCACTGCGCTTCTCTATTTTCAACATCTTTTACTGATCACATACAAGAT 1715  
 QY 1383 GCTATGCTCAGAGATATCCACAG 1407  
 Db 1716 CATTCGGCATGAAGATGTCCACAG 1740

RESULT 13  
 ABA00851  
 ID ABA00851 standard; DNA; 1386 BP.  
 XX  
 AC ABA00851;  
 XX  
 DT 16-APR-2003 (first entry)  
 XX  
 DE GLRA related nucleotide sequence #2.  
 XX  
 KW Glycine receptor subunit alpha; GLRA; p53 pathway; modulation; cancer;  
 breast; colon; lung; ovary; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200299140-A1.  
 XX  
 PD 12-DEC-2002.  
 XX  
 PF 02-JUN-2002; 2002WO-US017458.  
 XX  
 PR 05-JUN-2001; 2001US-0296076P.  
 PR 10-OCT-2001; 2001US-0328605P.  
 PR 15-FEB-2002; 2002US-0357253P.  
 XX  
 PA (EXEL-) EXELIXIS INC.  
 XX  
 PI Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;  
 WPI; 2003-167349/16.  
 XX  
 PT Identifying p53 pathway modulating agents, useful for the diagnosis and  
 treatment of disorders associated with defects in the p53 pathway, such  
 as cancer, comprises assaying glycine receptor subunit alpha (GLRA)  
 polypeptide or nucleic acid.  
 PT  
 XX  
 PS Disclosure; Page 41-42; 56pp; English.  
 CC The sequences given in ABA00850-55 represent glycine receptor subunit  
 alpha (GLRA) related nucleotide sequences. These sequences may be used in  
 the method of the invention for identifying a candidate p53 pathway  
 modulating agent. The method comprises: (a) assaying a purified GLRA  
 polypeptide or nucleic acid, or functionally active fragment or  
 derivative, with a test agent; (b) detecting a test agent-biased activity  
 of the assay system, and comparing the result with a reference. The  
 method is useful for identifying a candidate p53 pathway modulating  
 agent, modulating a p53 pathway of a cell and for diagnosing a disease in  
 a patient. Diagnosing a disease in a patient, comprising obtaining a  
 biological sample from the patient, contacting the sample with a probe  
 for GLRA expression, comparing results with a control, and determining  
 whether the results indicate a likelihood of disease. The disease is  
 cancer of the breast, colon, lung or ovary having a greater than 25%  
 expression level. The methods and compositions of the present invention  
 are useful for the diagnosis and treatment of disorders associated with  
 defects in the p53 pathway, such as cancer of the breast, colon, lung and  
 ovary  
 XX  
 SQ Sequence 1386 BP; 393 A; 320 C; 317 G; 356 T; 0 U; 0 Other;  
 Query Match 39.5%; Score 648.6; DB 10; Length 1386;  
 Best Local Similarity 70.0%; Pred. No. 5.3e-175;  
 Matches 942; Conservative 0; Mismatches 349; Indels 54; Gaps 3;  
 QY 63 CCTCTCAGGGTGGCTTGGCAAAAGAGGAAGTCAAATCTGGAACCAAGGGTCCAGCC 122  
 Db 68 CTTCAGGAGGGCTTCTGCAAAAGACCATGCTCCAGGTCTGGAACAACACCTTCACAGAC 127  
 QY 123 CATGTCCCCTCTGATTTCTAGACAAACTTATGGGGCGCAACATCTGGATATGATGCCAG 182  
 Db 128 CCTATCTCTTCAGATTTCTTGACAAAGTTTATGGGAAGGACATCAGGATATGATCGAAG 187  
 QY 183 GATTCGGCCCAATTTTAAAGGCCCCACCCCGTGAACCGTGAACCTTTCATCAACAG 242

188 AATCAGGCGCAAAATTTTAAAGGTCCTCCAGTAACGGTTACTTGCAATATTTTATCAACAG 247  
 243 TTTTACGTCGCTGACCAAGACCAATGACATACCGGGTGAATGCTCTTCTTGGGGCAACA 302  
 248 TTTTGGATCAGTACAGAAAGCAATGACATACCGAGTGAATATTTTCTGAGACACA 307  
 303 GTGGAATGACCCAGCGCTGCTTACCGAGAAATATCTGATGATCTCTTGAGACCTCGATCC 362  
 308 GTGGAATGATTCACGGCTGCGGTACAGTGAGTACCCAGATGACTCCCTGGACTTGGACCC 367  
 363 CTCATGCTGGACTCTATCTGAGAGCCAGACCTCTTCTTGTGTAATGAGAAAGGGGCCAA 422  
 368 ATCCATGCTAGACTCTCAATTTGGAACACAGATTTGTTCTTTGCGCAATGAGAAAGGGTCCAA 427  
 423 CTTTCCATGAGGTGACCAACGACCAAGTTTACTGCGCATCTTCAAGAAATGGGAATGCTGT 482  
 428 CTTTCCAGATGTACCCACTGACACAAATTTGCTACGGAATTTGAAATATGGCAAGTGTCT 487  
 483 GTACAGCATCAGGCTGACCCCTCATTTTGTCTGCTGATGAGCTCAAGAACTTCCCCAT 542  
 488 CTACAGTATCAGACTCACCTTGACCTTATCTGCTCCATGGACTTGAAGACTTTCCGAT 547  
 543 GGACATCCAGACGTGACAGATGAGCTTGAGAGCTCATCCATCTCTGAGCCCTCTGCC 602  
 548 GGATGTCCAGACCTGTACAAATGACGTGGAGATT----- 582  
 603 ATCTCTGTACATTTTCAATTTGGCTTACACCATGAAAGACCTCGTGTGAGTGGCTGAAGA 662  
 583 -----TTGGGTACAGATGAATGACCTGATATTTGAGTGGTTAAGTGA 625  
 663 TGCTCCTGTCTCAAGTGGCTGAGGGGCTGACTCTGCCCCAGTTTATCTTGGGATGA 722  
 626 TGGTC---CAGTCAAGTTGCTGAAGATTGACCTTGCCCGAGTTTATTTTGAAGAAGA 682  
 723 GAAGATCTAGGCTGTGTGACCAAGCACTACAAACAGGGAATTCACCTGATGATGAGT 782  
 683 GAAGGAATTTGGCTACTGTACAAAGCACTACAACTGGAAGATTTACCTGATGAGT 742  
 783 AAGTTTCACTTGAAGCGCAGATGGCTACTATCTGATTCAGATGATACATCCCGACCT 842  
 743 CAAATTTTCACTGGAAGCGCAATGGATATTTTGTGATTCAGATGATACATCCAGCT 802  
 843 ACTCATCGTATCTCTCTGCTGCTCTCTTCTGATCAACATGATGATGCTGCTGCTGCT 902  
 803 GCTTATAGTAAATTTTCTGCTGCTGCTCTTCTTGGATAAATATGATGATGATGCTGCT 862  
 903 TGTGGCGCTGGGATCACACCGCTGCTACCATGACACCCAGAGCTCTGCTGCTGCTGCT 962  
 863 GGTGCGACTGGGATCACACAGCTCTTAACGATGACACCCAGAGTTCAAGGCTCCAGGGC 922  
 963 CTCCTTGGCTTAAAGTGTCTTACTGTAAGCAATCGACATCGATGCTGCTGCTGCTGCT 1022  
 923 ATCTCTGCGCAAGGCTCTCTATGTAAGCGATTTGATCATCTGATGCGGCTGCTGCT 982  
 1023 CTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1082  
 983 GTTTGTGTTGCTGCTTACTGTAATACGAGGGTGAATCTGCTCTCAGGCAACACAA 1042  
 1083 AGAATTCATACGATTTGGAAGAGCAGAGCGGCCAACCGCTTGGAGGAAGATATCATCA 1142  
 1043 GGAGTTCCTGCGCTCCGGAAGAGACAGAGAGGCGAGATAAGGAAGAGAGCTTACTCG 1102  
 1143 AGAAGTCTCTTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1202  
 1103 TGAAGTCTGTTTAAATTTAGCGGTTATGGGATGGGTCTGCTGCTGCTGCTGCTGCTGCT 1162  
 1203 AGGTCCAAATGGAAGGTTCTGGCATTTATATGATCCCAACCTCCAGCCCTCTTCTAAGGA 1262  
 1163 AACAGCTGTCAAGGCCAC-----ACCTGCCAACCCACTCCCAACACCGCCCAAGA 1213  
 1263 AGAGAAACACCGCGGAATCTTACGTGAGCTGAGCCAGAGAAATTTGACCATCTCCCG 1322

1214 TGGAGATGCTATCAGAAAGAGTTTGTGACCGGGCAAAAGGATGACAGATATCTCG 1273  
 1323 GGTCTGTCTCCCTTTTACATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1382  
 1274 AGCTGCTTCCCATTTGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1333  
 1383 GCTATGCTCAGAAAGATATCCACAG 1407  
 1334 CATTCGGCATGAAGATGTCACAAG 1358

RESULT 14  
 ABA00850  
 ID ABA00850 standard; DNA; 1854 BP.  
 XX ABA00850;  
 AC ABA00850;  
 XX 16-APR-2003 (first entry)  
 DT XX  
 DE GLRA related nucleotide sequence #1.  
 XX  
 KW Glycine receptor subunit alpha; GLRA; p53 pathway; modulation; cancer;  
 KW breast; colon; lung; ovary; ss.  
 OS Homo sapiens.  
 XX  
 XX WO200299140-A1.  
 PN 12-DEC-2002.  
 PD 02-JUN-2002; 2002WO-US017458.  
 XX 05-JUN-2001; 2001US-0296076P.  
 PR 10-OCT-2001; 2001US-0328605P.  
 PR 15-FEB-2002; 2002US-0352553P.  
 XX (EXEL-) EXELIXIS INC.  
 PI Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;  
 XX WPI; 2003-167349/16.  
 DR  
 XX Identifying p53 pathway modulating agents, useful for the diagnosis and  
 PT treatment of disorders associated with defects in the p53 pathway, such  
 PT as cancer, comprises assaying glycine receptor subunit alpha (GLRA)  
 PT polypeptide or nucleic acid.  
 XX Example; Page 40-41; 56pp; English.

The sequences given in ABA00850-55 represent glycine receptor subunit  
 alpha (GLRA) related nucleotide sequences. These sequences may be used in  
 the method of the invention for identifying a candidate p53 pathway  
 modulating agent. The method comprises: (a) assaying a purified GLRA  
 polypeptide or nucleic acid, or functionally active fragment or  
 derivative, with a test agent; (b) detecting a test agent-biased activity  
 of the assay system, and comparing the result with a reference. The  
 method is useful for identifying a candidate p53 pathway modulating  
 agent, modulating a p53 pathway of a cell and for diagnosing a disease in  
 a patient. Diagnosing a disease in a patient, comprising obtaining a  
 biological sample from the patient, contacting the sample with a probe  
 for GLRA expression, comparing results with a control, and determining  
 whether the results indicate a likelihood of disease. The disease is  
 cancer of the breast, colon, lung or ovary having a greater than 25%  
 expression level. The methods and compositions of the present invention  
 are useful for the diagnosis and treatment of disorders associated with  
 defects in the p53 pathway, such as cancer of the breast, colon, lung and  
 ovary

Sequence 1854 BP; 525 A; 415 C; 416 G; 498 T; 0 U; 0 Other;  
 Query Match 39.5%; Score 648.6; DB 10; Length 1854;  
 Best Local Similarity 70.0%; Pred. No. 6.2e-175;  
 Matches 942; Conservative 0; Mismatches 349; Indels 54; Gaps 3;



QY 63 CCTCCTCAGGTGGCTTGGCAAGAGAGAGTCAATCTGAAACAAGGGGTCCAGCC 122  
 Db 447 CTTACAGGCGGCTTCTGCAAGAGACCATGCTCCAGGTCTGGAACAACAACCTTTCAGAC 506  
 QY 123 CATGTCCTCTGATTTCTAGACAACTTATGGGGCGAATCTGGATATGATGCCAG 182  
 Db 507 CTTATCTCTCTGATTTCTGGACAACTTATGGGAAGACATCAGGATATGATGCCAG 566  
 QY 183 GATTCGGCCCAATTTTAAAGGCCCAACCCGTGAACGTGACCTGCAACATCTTTCATCAACAG 242  
 Db 567 AATCAGGCCAAATTTTAAAGGCTCTCCAGTAAAGGTTACTTGCATATTTTATCAACAG 626  
 QY 243 TTTACGTCCTGTCACCAAGACCAATGACCTACCGGGTGAATGCTTCTTGGGGCAACA 302  
 Db 627 TTTTGGATCAGTCACAGAAACCAATGACCTACCGAGTGAATATTTTCTGAGACAACA 686  
 QY 303 GTTGAATGACCAAGCGCTGCTTACCGAATATCTGATGACTCTCTGCACTCGATCC 362  
 Db 687 GTTGAATGATTCACGGCTGGCTGACGATGAGTACCCAGATGACTCTCTGCACTTGGACCC 746  
 QY 363 CTCATGCTGGACTCTATCTGGAAGCCAGACCTCTTCTTTGCTTAATGAGAAAGGGCCAA 422  
 Db 747 ATCCATGCTAGACTCCATTTGGAACACAGATTTGTTCTTTGCAATGAGAGGGTCCAA 806  
 QY 423 CTTCCATGAGGTGACCAACGCAACAAAGTTATCTGCGCATCTTCAAGATGGGAATGTCT 482  
 Db 807 CTTCCACGATGTCAACCACTGACAACTTGTACGGATTTGGAATAATGGCAAGTGTCT 866  
 QY 483 GTACAGCATCAGCTGACCTCATTTTGTCTGCTGATGGACCTCAAGAACTTCCCAT 542  
 Db 867 CTACGATACAGACTCAGCTTACCTTACCTTCTGCTCCAGTCTTGAAGAACTTTCCGAT 926  
 QY 543 GGACATCCAGAGCTGACGATGACGCTTGGAGCTCATCATCTCTGCAAGCCCTCTGCC 602  
 Db 927 GGATGTCCAGACTGTACATGACGCTGGAGTT----- 961  
 QY 603 ATCTCTGTCATTTGATTTGGCTGTACACCATGAAAGACCTCGTGTGAGTGGCTGGAAGA 662  
 Db 962 -----TTGGGTACACGATGAATGACCTGATATTTGAGTGGTAAAGTGA 1004  
 QY 663 TGCTCTCTGCTGCAAGTGGCTGAGGGCTGACTCTGCCCCAGTTTATCTTGGGGATCA 722  
 Db 1005 TGGTC---CAGTGAAGTTGCTGAAGGATGACCTCGCCCCAGTTTATTTTGAAGAAGA 1061  
 QY 723 GAAGGATCTAGGCTGTGTACCAAGACCTACACACAGGGAATTCACCTGATCGAGGT 782  
 Db 1062 GAAGGAACTTGGCTACTGTACAAAGCACTACACACTGGAAGTTTACCTGCAATTGAGT 1121  
 QY 783 AAAGTTTCACTGGAAACGGCAGATGGGCTACTATCTGATTCAGATGATCATCCCCAGCCT 842  
 Db 1122 CAAAGTTTCACTGGAAACGGCAGATGGGCTACTATTTGATCCAGATGATCATCCCCAGCCT 1181  
 QY 843 ACTCATCTGCTGCTCTCTGCTGCTCTCTCTGATCAACATGGATGCTGCCCTGCCCG 902  
 Db 1182 GCTTATAGTAATTTCTGCTGCTGCTCTCTTTTGGATAAATATGATGCAAGCCCTCCAG 1241  
 QY 903 TGTGGCCCTGGGATCACCACCGTGTCTACCATGACACCCAGCTCTGGCTCCCGGC 962  
 Db 1242 GGTGCACTTGGGATCACCACAGTCTTACGATGACACCCAGGTTGAGGCTCAGGGC 1301  
 QY 963 CTTCTTGGCTTAAAGGTCTCTACGTGAAGGCAATCGACATCTGGAATGGCTGTGTCTGCT 1022  
 Db 1302 ATCTCTGCAAAAGGTCTCTATGTAAAGGATTTGACATCTGATGGCGGTGTGCTTCT 1361  
 QY 1023 CTTTGTGCTGCTGCTCTGCTGAGTATGCTGCAATAAATTTTGTCTGCTGAGATTA 1082  
 Db 1362 GTTTGTGTTGCTGCTTACTGGAATACCGAGCGGTGAACTTCTCTCCAGGCAACAA 1421  
 QY 1083 AGAATTCATACGATTCGGAAGAGGAGGAGGCGCCCAACGCTTGGAGGAAGATATATCA 1142  
 Db 1422 GGAGTTCTCTGGCTCCGGAAGAAGACAGAGAGGCAAGATAGGAAGAAGACGTTACTCG 1481

QY 1143 AGAAGTCTGTTTCTATTTCCGTTGGCTATGGCTTGGSCCACTGCTGCAAGAGATGG 1202  
 Db 1482 TGAAGTCTGTTTAAATTTTACGGTTATGGGATGGTCTACTGCTTCAAGTGAAGATGG 1541  
 QY 1203 AGGTCCAATGGAAGGTCTCTGGCATTTATAGTCCCAACCTCCAGCCCTCTTCTTAAGGGA 1262  
 Db 1542 AACAGTCTCAAGGCCAC-----ACCTGCCAACCCACTCCCAACACCGCAAAAGA 1592  
 QY 1263 AGAGAAACCAACCGGAAACTCTACGTGGAATGAGCAAGAGAAATTTGACACCATCTCCCG 1322  
 Db 1593 TGGAGATGCTATCAAGAAGAAAGTTTGTGACCGGGCAAAAGGATTTGACACCATATCTCG 1652  
 QY 1323 GCCTGCTTCCCTTTTCACTTCTCTCATCTTCAATATCTTCTACTGGTGTCTATAAAGT 1382  
 Db 1653 AGCTGCTTCCCACTTGGCCTTCTCATTTTCAACATCTTTTACTGGATCACAATCAAGAT 1712  
 QY 1383 GCTATGGTCAGAAGATATCCACAG 1407  
 Db 1713 CATTGGCATGAAGATGTCACAG 1737

RESULT 15  
 ADC77686  
 ID ADC77686 standard; cDNA; 3069 BP.  
 XX AC ADC77686;  
 XX DT 01-JAN-2004 (first entry)  
 XX Human 2387 protein encoding cDNA SEQ ID NO:49.  
 DE pain disorder; pain signalling mechanism; analgesic; antimigraine;  
 KW antinflammatory; gene therapy; inflammatory pain; chronic pain;  
 KW neuropathic pain; neuralgia; fibromyalgia; cancer pain; migraine;  
 KW headache; pain; human; gene; ss.  
 XX Homo sapiens.  
 OS Key  
 FH Location/Qualifiers  
 CDS 421..1770  
 FT /\*tag= a  
 FT /product= "2387 protein"  
 XX PN WO2003073983-A2.  
 XX 12-SEP-2003.  
 XX 19-FEB-2003; 2003WO-US004816.  
 XX 28-FEB-2002; 2002US-0360495P.  
 PR 04-APR-2002; 2002US-0370121P.  
 PR 16-APR-2002; 2002US-0373010P.  
 PR 19-APR-2002; 2002US-0373908P.  
 PR 03-MAY-2002; 2002US-0377717P.  
 PR 13-MAY-2002; 2002US-0379949P.  
 PR 21-MAY-2002; 2002US-0382409P.  
 PR 06-JUN-2002; 2002US-0386879P.  
 PR 10-JUN-2002; 2002US-0387536P.  
 PR 08-JUL-2002; 2002US-0394376P.  
 PR 21-AUG-2002; 2002US-0404996P.  
 PR 09-SEP-2002; 2002US-0412006P.  
 PR 10-OCT-2002; 2002US-0417327P.  
 PR 15-NOV-2002; 2002US-0417499P.  
 PR 10-DEC-2002; 2002US-0426964P.  
 XX PA (MILL-) MILLENNIUM PHARM INC.  
 XX Rosenfeld JB, Silos-Santiago I;  
 XX WPI; 2003-712843/67.  
 DR P-PSDB; ADC77687.

XX Identifying a compound capable of treating a pain disorder e.g.,  
PT neuropathic pain comprises assaying the ability of the compound to  
PT modulate the nucleic acid expression or polypeptide activity.  
PS Claim 1; SEQ ID NO 49; 277bp; English.  
XX  
CC The present invention describes a method for identifying a compound (C)  
CC capable of treating a pain disorder comprising assaying the ability of  
CC the compound to modulate 9949, 14230, 760, 62553, 12216, 17719, 41897,  
CC 47174, 33408, 10002, 16209, 314, 636, 27410, 33260, 619, or 13424 nucleic  
CC acid expression or 9949, 14230, 760, 62553, 12216, 17719, 41897, 47174,  
CC 33408, 10002, 16209, 314, 636, 27410, 33260, 619, 15985, polypeptide  
CC activity. Also described: (1) identifying a compound (C) capable of  
CC modulating a pain signalling mechanism; and (2) treating a subject having  
CC a pain disorder characterised by aberrant nucleic acid expression or  
CC polypeptide activity. (C) has analgesic, antimigraine and  
CC antiinflammatory activities, and can be used in gene therapy. The method  
CC is useful for identifying a modulator compound capable of treating a pain  
CC disorder, e.g. inflammatory pain, chronic pain, neuropathic pain,  
CC neuralgia, fibromyalgia, cancer pain, migraine/headache pain or tissue  
CC pain comprising administering the modulator to a subject having a pain  
CC disorder characterised by aberrant nucleic acid expression or polypeptide  
CC activity. The present sequence encodes the human 2387 protein from the  
CC present invention.  
XX  
XX Sequence 3069 BP; 962 A; 547 C; 583 G; 977 T; 0 U; 0 Other;  
Query Match 33.7%; Score 552.8; DB 10; Length 3069;  
Best Local Similarity 66.6%; Pred. No. 2.5e-147;  
Matches 885; Conservative 0; Mismatches 377; Indels 66; Gaps 4;  
QY 80 TGSCAAAGAGGAAGTCAATCTGGAAACCAAGGGGTCCAGCCCATGTCCCTCTGATT 139  
DB 497 TTGCCCAAGAAACAGACAGTGCAGATCTCGAAGTCTCAATGTTCACCTTCTGATT 556  
QY 140 TCCTAGACAACTTATGGGGCGAACAATCTGGATATGATCGAGATTTCGGCCCAATTTTA 199  
DB 557 TTCTGGATAAATTAATGGGCGAGACATCAGGATATGATCGAAGATCAGACCAATTTTA 616  
QY 200 AGGCCCAACCGTGAACGTGACCTGCAACATTTTCATCAACAGTTTCAGCTCGGTCAACA 259  
DB 617 AAGGCCCTCCAGTTAATGTACATGCAACATATTCATCAACAGTTTCGGCTCTATCGCAG 676  
QY 260 AGACCAACATGGAACCTCGGTCGAATGCTCTTGGGCAACAGTGGATGACCCACGCC 319  
DB 677 AGACGACCATGGATTACAGATGATATCTTCTTCTGCAAGATGGAATGATCCCCGCC 736  
QY 320 TGTCTTACCGAGAAATCTCTGATGACTCTCTGGACCTCGATCCCTCCATCTGGACTCTA 379  
DB 737 TCGCGTACAGTGAATATCTGACGACTCTTTAGACCTCGACCCCTCCATGTTGGACTCCA 796  
QY 380 TCTGGAAGCCAGACCTCTTCTTGTCTAATGAGAAAGGGGCAACTTCCATGAGGTGACCA 439  
DB 797 TTTGGAAACCTGATTTGTTCTTTGGCAATGAAAGGGTGCACACTTTCATGAAGTCACTA 856  
QY 440 CGGACCAACAAAGTTACTCGGCATCTTCAAGATGGGAATGCTGTACAGCATCAGGCTGA 499  
DB 857 CAGACCAACAAATGCTAAGAAATTTTCAAAATGGAATGTTCTTTAATCAATGAATTA 916  
QY 500 CCCTCAATTTTGTCTCGCTGATGGACCTCAAGAACTTCCCCATGGACATCCAGACGTGCA 559  
DB 917 CATTAACACTTTCCTGCAATGATCTCAAGAAATTTTCCCATGGATGTACAAACATGTA 976  
QY 560 CGATGAGCTTGAAGCTCATTCATCTCTGACGCCCTCTGCCATCTCTGTCACTTTTCAG 619  
DB 977 TAATGCAACTGGAAAGCT----- 994  
QY 620 TTGGCTACACCATGAAGACCTCGTGTGTTGAGTGGCTGGAAGATGCTCCTGCTGCCAAG 679  
DB 995 TTGGGTACACATGAATGATCTCATTTTGAATGGC---AAGATAGGCCACCGGTACAAG 1051  
QY 680 TGGCTGAGGGGCTGACTCTGCCCCAGTTTATCTTTGGGGATGAGAGGATCTAGGCTGTT 739

1052 TGGCAGAAGGACTCACTTTGCCCCAGTTTCTGTTGAAAGAGAAAAAGATTTACGATACT 1111  
QY 740 GTACCAAGCACTACACACAGAGGAATTCACCTGCATCGAGGTAAAGTTTTCACCTGGAAC 799  
DB 1112 GCACATAACATTACAAATACAGGAAAGTTTACGTGTATAGAGTGGGATTCCTATCTGGAGC 1171  
QY 800 GGCAGATGGGCTACTATCTGATTCAGATGATACATCCCAAGCCTACTCATCGTCTATCTGT 859  
DB 1172 GACAAATGGGATACTATCTGATCCAGATGATACATTCCTAGTCTCTGATGTTTATCTAT 1231  
QY 860 CTGGGTCTCTCTTGGATCAACATGGATGTCGCCCTGCGCTGCGCTTGGCCCTGGGCATCA 919  
DB 1232 CTGGGTCTCTCTTGGATCAACATGGATGTCAGCACCGGCCAGGGTAGCTCTGGGGATAA 1291  
QY 920 CCACGCTGCTCACCATGACCAACCAAGAGCTCTGGCTCCCGGGCCTCTTGGCCTAAGGTGT 979  
DB 1292 CCACCGTGTAAACGATGACTACACAGAGTTACAGAGCTTCCTTTGCCAAAGATT 1351  
QY 980 CTACGTGAAGCAATCGACATCTGATGGCTGTGTCTGCTCTTGTGTTGCTGCTGCT 1039  
DB 1352 CATATGTCAAAGCTATTGATATTTGGATGGCAGTATGCTCTCTTGTGTTGTTTTCAGCAC 1411  
QY 1040 TCGTGGATGATCTGCCATAAATTTTGTCTCGTCAGCATAAAGAAATTCATACGACTTC 1099  
DB 1412 TTCTGGAGTATGCAGCTGTAAATTTGTATCAAGACACACAAAGAACTTCTGAGATTTC 1471  
QY 1100 GAAGAGGACAGAGCGGCCAACCGTTGGAGGAGATATCATCCNAGAAAGTCTGTTCTATT 1159  
DB 1472 GACGAAAGAGAAAG-----AATAAGGATGATGAGGTAAAGGAAAGCCGATTCAGCT 1522  
QY 1160 TCGTGGCTTATGGCTTGGGCCACTGCTGCAGGCAAGAGATGGAGGTCCAATGGAAAGTT 1219  
DB 1523 TCACAGCTTATGGATGGGACCATGCTTACAGCAAGAGATGGCATGACTCCAAAGGGCC 1582  
QY 1220 CTGGCATTTATAGTCCCCAACCTCCAGCCCTCTTCTAAGGGAAGAGAGAAACACCGCGGA 1279  
DB 1583 C-----CAACCACTGTCCAGGTAAATGCCAAAAAGTCTGATGAAATGAGGA 1630  
QY 1280 AACTCTACGTGACCTGAGCCAGAGAAATTCACCATCTCCCGGCTGTCTTCCCTTTCA 1339  
DB 1631 AGGTCTTTATCGACCGGGCCAAAGAGATTGATACCATCTCCGAGCCTCTTCCCAATTAG 1690  
QY 1340 CTTTCTCATCTTCAATATCTTCTACTGGTGTCTATAAAGTGTATGCTCAGAGAGATA 1399  
DB 1691 CTTTGTGATTTTAAATATTTTCTACTGGTTATCTATAAAATTTCTTAGGCATGAGGATA 1750  
QY 1400 TCCACGAG 1407  
DB 1751 TTCTATCAG 1758

Search completed: April 28, 2005, 12:05:58  
Job time : 1101.49 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 28, 2005, 11:47:18 ; Search time 337.133 Seconds  
(without alignments)  
7959.755 Million cell updates/sec

Title: US-10-075-846-3  
Perfect score: 1640  
Sequence: 1 agacaactctgtctctgc.....taaaaaaaaaaaaaaaaaa 1640

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA: \*  
1: /cgn2\_6/prodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/prodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/prodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/prodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/prodata/1/ina/PCTUS\_COMB.seq.\*  
6: /cgn2\_6/prodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	656	40.0	1372	4	US-09-949-016-4602 Sequence 4602, Ap
2	656	40.0	1715	4	US-09-949-016-27 Sequence 27, Appl
3	648.6	39.5	1854	4	US-09-949-016-5267 Sequence 5267, Ap
4	253	15.4	253	4	US-09-016-434-271 Sequence 271, Appl
5	211	12.9	2109	4	US-09-949-016-3698 Sequence 3698, Ap
6	193	11.8	1555	3	US-08-809-802-11 Sequence 11, Appl
7	187	11.4	4621	3	US-09-592-891A-13 Sequence 13, Appl
8	187	11.4	4621	4	US-09-969-844-13 Sequence 13, Appl
9	182.2	11.1	2310	1	US-08-417-330A-15 Sequence 15, Appl
10	181.6	11.1	1609	3	US-09-592-891A-11 Sequence 11, Appl
11	181.6	11.1	1609	4	US-09-969-844-11 Sequence 11, Appl
12	179	10.9	3958	1	US-08-435-933-5 Sequence 5, Appl
13	179	10.9	3958	5	PCT-US96-06035-5 Sequence 5, Appl
14	176.8	10.8	73519	4	US-09-949-016-16344 Sequence 16344, A
15	176.8	10.8	105919	4	US-09-949-016-11769 Sequence 11769, A
16	174.2	10.6	1640	3	US-09-592-891A-12 Sequence 12, Appl
17	174.2	10.6	1640	4	US-09-969-844-12 Sequence 12, Appl
18	168	10.2	1368	3	US-09-130-339-1 Sequence 1, Appl
19	161.2	9.8	1297	3	US-09-130-339-3 Sequence 3, Appl
20	158.8	9.7	1884	1	US-09-403-667A-3 Sequence 3, Appl
21	157.8	9.6	1657	1	US-08-554-659-3 Sequence 3, Appl
22	157.8	9.6	1844	3	US-09-002-361-1 Sequence 1, Appl
23	157.2	9.6	1884	4	US-09-403-667A-1 Sequence 1, Appl
24	156.4	9.5	1866	1	US-08-417-330A-19 Sequence 19, Appl
25	156.2	9.5	1657	1	US-08-554-659-1 Sequence 1, Appl
26	156.2	9.5	1732	1	US-08-417-330A-17 Sequence 17, Appl
27	155	9.5	205163	4	US-09-949-016-17009 Sequence 17009, A

28 146.6 8.9 669 3 US-09-002-361-7 Sequence 7, Appl  
29 143.2 8.7 1997 4 US-09-620-312D-256 Sequence 256, App  
30 141.8 8.6 1491 1 US-08-137-614A-3 Sequence 3, Appl  
31 141.8 8.6 1491 2 US-08-768-301-1 Sequence 1, Appl  
32 138.6 8.5 1638 1 US-08-417-330A-13 Sequence 13, Appl  
33 137.6 8.4 1519 3 US-09-002-361-4 Sequence 4, Appl  
34 136.2 8.3 1038 4 US-09-765-069-9 Sequence 9, Appl  
35 136.2 8.3 1179 4 US-09-765-069-3 Sequence 3, Appl  
36 136.2 8.3 1263 4 US-09-765-069-7 Sequence 7, Appl  
37 136.2 8.3 1404 4 US-09-765-069-1 Sequence 1, Appl  
38 136.2 8.3 1987 1 US-09-620-312D-331 Sequence 331, App  
39 135.6 8.3 2191 1 US-08-417-330A-11 Sequence 11, Appl  
40 135.2 8.2 1052 3 US-09-592-891A-8 Sequence 8, Appl  
41 135.2 8.2 1052 4 US-09-969-844-8 Sequence 8, Appl  
42 131.6 8.0 2066 3 US-08-072-064-7 Sequence 7, Appl  
43 130.6 8.0 1693 4 US-09-620-312D-131 Sequence 131, App  
44 130 7.9 2066 3 US-08-072-064-2 Sequence 2, Appl  
45 128.8 7.9 1679 4 US-09-016-434-1470 Sequence 1470, Ap

## ALIGNMENTS

## RESULT 1

US-09-949-016-4602  
; Sequence 4602, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CLO01307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4602  
; LENGTH: 1372  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-4602

Query Match 40.0%; Score 656; DB 4; Length 1372;  
Best Local Similarity 71.9%; Pred. No. 1.1e-186;  
Matches 952; Conservative 0; Mismatches 315; Indels 57; Gaps 5;  
QY 89 AGGAAGTCAAATCTGGAAACCAAGGGGTCCAGGCCATGTCCTCTGATTTCCTAGACA 148  
DB 26 AGAGGCTGAGCTGCTCGTCCGACCAAGCCTATGTCACCTCGGATTCCTGGATA 85  
QY 149 AACTTATGGGCGAACAATCTGGATATGATGCCAGGATTCGGCCCAATTTAAAGGCCAC 208  
DB 86 AGCTAATGGGAGAACCTCCGGATATGATGCCAGGATTCAGGCCCAATTTAAAGGTCCTC 145  
QY 209 CGGTGAACGTGACCTGCAACATCTTCATCAACAGTTTACGTCCTCGTCACCAAGACCACAA 268  
DB 146 CAGTGAACGTGAGCTGCAACATTTTCATCAACAGCTTTGGTTTCATTTGCTGAGACACCA 205  
QY 269 TGGACTACCGGGTGAATGCTCTTCGCGCAACAGTGGGAATGACCCAGCGCTGCTCTACC 328  
DB 206 TGGACTATAGGTCAACATCTTCCTCGGAGCAATGGAGACGCCCGCTGGCTTATA 265  
QY 329 GAGAAATATCTGATGACTCTCTGGACCTCGATTCCTCCATGCTGGACTCTATCTGGAAGC 388  
DB 266 ATGAATACCTGACGACTCTCTGGACCTGGACCCATCCATGCTGGACTCTCATCTGGAAC 325  
QY 389 CAGACCTCTTTTGTGTAATGAGAAAGGGGCCAACTTCCATGAGGTGACCCAGGACACAA 448

```

Db      326 CTGACCTGTTCTTGGCAACGAGAGGGGCCACCTTCCATGAGATCACCACAGACAAC 385
QY      449 AGTTACTCGGCATCTTCAAGATGGGAATGTGCTGTACAGCATCAGGCTGACCTCATTT 508
Db      386 AATTGCTAAGGATCTCCCGAATGGGAATGTCTCTACAGCATCAAAATCACCCTGACAC 445
QY      509 TGTCTGCTGATGGACCTCAAGAACTCCCATGACATCCAGAGCTGCAGATGCAGC 568
Db      446 TGGCTTGGCCATGGACTTGAAGATTTCCCATGAGATGTCCAGATGATATCATGCAAC 505
QY      569 TTGAGAGCTCATCATPACTCTGAGCCCTCTGCAATCTCTGTCACTTTCCAGTTGGCTACA 628
Db      506 TGGAAAGCT-----TTGGATATA 523
QY      629 CCATGAAGACCTCGTGTGTTGAGTGGCTGGAAGATGCTCCTGTCTCCAGTGGCTGAGG 588
Db      524 CGATGAATGACCTCATCTTTGAGTGGCAGGAACAGG---GAGCGGTGACAGGATG 580
QY      689 GGCTGACTCTGCCCAAGTTTATCTTCGGGATGAGAGGATCTAGGCTGTGTACCAAGC 748
Db      581 GACTAATCTGCCCAAGTTTATCTTCGAAGAGAGAGGACTTGAATGATGATGACCAAGC 640
QY      749 ACTACAACACAGGGAATTCACCTGCATCGAGGTAAAGTTTCACTTGGAAACGCGAGATGG 808
Db      641 ACTACAACACAGGTAATTCACCTGCATTTGAGGCCCGGTTCCACCTGGAGCGCAGATGG 700
QY      809 GCTACTATCTGATTCAGATGTATATCCCGAGCTTACTCATCTGCTATCTCTGCTGGTCT 868
Db      701 GTTACTACCTGATTCAGATGTATATCCCGAGCTGCTCATTTGTCATCTCTCATGATCT 760
QY      869 CTTCTGGATCAACATGGATGCTGCCCTGCCGTGGGCTGGGCTGAGCATTACCAAGC 928
Db      761 CTTCTGGATCAACATGGATGCTGACCTGCTGCTGGGCTGAGGCTGAGCATTACCAAGC 820
QY      929 TCACATGACCAACCCAGAGCTCTGGCTCCCGGCTCTTTTGGCTAAGGTGCTCTACGTGA 988
Db      821 TCACATGACCAACCCAGAGCTCTGGCTCTCGAGCATCTCTGCCCAAGGTGCTCTATGTC 880
QY      989 AGCAATGACATCTGGATGGCTGTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1048
Db      881 AAGCCATTGACATTTGGATGGCAGTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 940
QY      1049 ATGCTGCCAATAATTTGTTCTGCTGAGCAAGAAATTCATACGACTTCGAAGAGGC 1108
Db      941 ATGCTGCCGTTAATTTGTTCTGCTGCAACATTAAGAGCTGCTCCGATTCAGAGGAAGC 1000
QY      1109 AGAGCGCCAAACGCTTGGAGGAAGATATATCAACAGAAAGTGTGTTTCTATTTCCGTTGCT 1168
Db      1001 GGAGACATCAAGAGGAGTGAAGCT-----GGAGAGGCGCTTTAATCTCTGCTGCT 1054
QY      1169 ATGGCTTGGCC---ACTGCTGCGAGGAGATGAGGTCCAATGGAGGTTG---TG 1222
Db      1055 ATGGGATGGGCCAGCTGCTTACAGGCCAAGGATGGCATCTCAGTCAAGGGCGCCAA 1114
QY      1223 GCATTTATAGTCCCAACCTCCAGCCCTCTTCTAAGGGAAGAGAAACACAGCGGAAC 1282
Db      1115 ACATGAACCAACACCCCTCTCTGACATCTAAGTCCCAAGGAGATGCGAAGAC 1174
QY      1283 TCTACGTGGAATGAGCAAGAGAAATGACACCATCTCCCGGGCTGTCTTCCCTTTCACTT 1342
Db      1175 TCTTCATCAGAGGGCCCAAGAGATCGACAAATATCCCGCATTTGGCTTCCCGATGGCCT 1234
QY      1343 TCCATCTTCAATATCTTCTACTGGGTGTCTATAAAGTGTATGGTCAGAGATATCC 1402
Db      1235 TCCATCTTTCAACATGTTCTACTGATCATCTACAAGATTTGTCGCTAGAGAGGAGCTCC 1294
QY      1403 ACCA 1406
Db      1295 ACNA 1298

```

```

US-09-949-016-27
; Sequence 27, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 1715
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-27

Query Match      40.0%; Score 656; DB 4; Length 1715;
Best Local Similarity 71.9%; Pred. No. 1.2e-186;
Matches 952; Conservative 0; Mismatches 315; Indels 57; Gaps 5;

```

```

QY      89 AGGAAGTCAATCTCGAACCAAGGGTCCAGCCCATGTCCCCCTCTGATTTCTTAGACA 148
Db      367 AGGAGGCTGAAGCTGCTCGTCCGCAACCAAGCCATGTACCCCTCGGATTTCTTGATA 426
QY      149 AACTTATGGGCGAATCTCTGATATGATGCGAGATTCGGCCCAATTTTAAAGGCCAC 208
Db      427 AGCTAATGGGAGAACCTCCGGATATGATGCGAGATTCAGGCCCAATTTTAAAGGTCCC 486
QY      209 CCGTGAAGCTGACCTGCAACATCTTTCATCAACAGTTTTCAGCTCCGTCACCAAGACCA 268
Db      487 CAGTGAACGTGAGCTGCAACATTTTCATCAACAGCTTTGTTCCATTTGCTGAGACA 546
QY      269 TGGACTACCGGTGAATGTCTTCTTGGGCAACAGTGGATGACCAAGCCCTGCTTAC 328
Db      547 TGGACTATAGGTCAACATCTTCTCGGCGAGCAATGGAAACGACCCCGCTGGCTATA 606
QY      329 GAGAATATCTGATGATCTCTGACCTCTGAGCTTCATGCTCCATGCTGACTCTATCTGAAGC 388
Db      607 ATGAATACCTGACGACTCTCTGACCTGAGACCCATTCATGCTGACTCCATCTGGAAC 566
QY      389 CAGACCTCTTCTTGTCTAATGAGAAAGGGGCCAATCTCCATGAGGTGACACGGAACA 448
Db      667 CTGACCTGTTCTTGGCAACGAGAGGGGGCCCACTTCCATGAGATCACCACAGACA 726
QY      449 AGTTACTGGCATCTTCAAGATGGGAATGTCTGTACAGCATCAGGCTGACCTCATTT 508
Db      727 AATTGCTAAGGATCTCCCGAATGGGAATGTCTCTACAGCATCAGAAATCACCCTGACAC 786
QY      509 TGTCTGCTGATGAGCCTCAAGAACTTCCCATGACATCCAGAGCTGCAGATGCAGCAG 568
Db      787 TGGCTTGGCCCATGAGACTTGAAGATTTCCCATGATGTCCAGATGATATCATGCAAC 846
QY      569 TTGAGAGCTCATCCATACTCTGACGCCCTCTGCCCATCTCTGTCACTTTTCAGTTGGCTACA 628
Db      847 TGGAAAGCT-----TTGGATATA 864
QY      629 CCATGAAGACCTCGTGTGTTGAGTGGCTGGAAGATGCTCCTGCTGCCAAGTGGCTGAGG 688
Db      865 CGATGAATGACCTCATCTTGGAGTGGCAGGAACAGG---GAGCGGTGACAGGATG 921
QY      689 GGCTGACTCTGCCAGTTTATCTTGGGATGAGAGGATCTAGGCTGTGTGACCAAGC 748
Db      922 GACTAATCTGCCCCAGTTTATCTTGAAGAGAGAGGAGCTTGAATGATGATGATGATGATG 981
QY      749 ACTACAACACAGGGAATTCACCTGATCGAGGTAAAGTTTTCACCTGGAACCGCAGATGG 808

```







```

QY 462 CTTCAAGATGGGAATGCTGTACAGCATCAGGCTGACCTCATTTTGTCTGCTGAT 521
Db 573 TTTTGTGATGGAGATGCTCTGTACAGCATGAGGTATCTATTTATCTTTTCATGCCCCTTT 632
QY 522 GGACCTCAAGAACTTCCCATGACATCCAGAGCTGCAGCATGAGCTTGAGAGCTCATC 581
Db 633 GGACTTGACATTTTCCCATGATACACAACTTGCAGATGCACTGGAGACT--- 688
QY 582 CATACTCTGAGCCCTCTGCTCATCTGTCACTTTCAGTTGGCTACACATGAAGACCT 641
Db 689 -----TTGGTTACAACTGATGATTT 710
QY 642 CGTGTGTTGAGTGGCTGGAAGATCTCTGCTGTCCAAGTGGCTGAGGGCTGACTCTGCC 701
Db 711 ACATTTATCTGGCAGTCCAGGATCTCTGTGCAATTAGAAAAATTTGCCTTGCCTCAAT 770
QY 702 CCAGTTTATCTTCGGGATGAGAAGATCTAGGCTGTGTACCAAGCACTTACAAACAGG 761
Db 771 TGATATCAAAAAGAGATATTGAATATGTTAATCTGTACAAATACTATAAGGCAAGG 830
QY 762 GAAATTCACCTGATCGAGTAAGTTTCACTTGAACGAGATGGGCTACTATCTGAT 821
Db 831 CTACTACACATGCTGGAAGTCTATCTTCACTTGAAGGAGGCTGCGCTTTTACATGAT 890
QY 822 TCAGATGTATATCCAGCTTACTCTGCTCATCTCTGCTGCTGCTGCTGCTGCTGCTGCT 881
Db 891 GGGGGTCTAGCCCAACCTCTCATTTGTTGTTCTCTCTGCTTCTGCTTCTGCTGCTGCT 950
QY 882 CATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 941
Db 951 CCGGAGCGGAGTGTGCGAGATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1010
QY 942 CCAGAGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1001
Db 1011 TGAGTGACAACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1070
QY 1002 CTGGATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1054
Db 1071 TTGGCTTATTGCTTGGCTTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1123

```

RESULT 6

```

US-08-809-802-11
; Sequence 11, Application US/0809802
; Patent No. 6455276
; GENERAL INFORMATION:
; APPLICANT: Le Bourdelles, Beatrice
; APPLICANT: Whiting, Paul John
; TITLE OF INVENTION: HUMAN ALPHA 4 RECEPTOR SUBUNIT
; TITLE OF INVENTION: OF THE GABA-A RECEPTOR
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FASTSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/809,802
; FILING DATE: 19-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/02323
; FILING DATE: 29-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Yablonsky, Michael D
; REGISTRATION NUMBER: 40,407

```

```

; REFERENCE/DOCKET NUMBER: T1292
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-4678
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1555 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 47...1402
; OTHER INFORMATION:
; US-08-809-802-11

Query Match 11.8%; Score 193; DB 3; Length 1555;
Best Local Similarity 53.3%; Pred. No. 2.9e-47;
Matches 486; Conservative 0; Mismatches 380; Indels 45; Gaps 2;

QY 142 CTAGACAAACTTTATGGGGGCAACATCTGGATATGATCCAGGATTCGGCCCAATTTTAA 201
Db 167 CTCCTCCAACTGGACGGGCTGATAGCGGTACGCCGCAACTTCGGGCTGGCATCGGA 226
QY 202 GGGCCCAACCGTGAACTGACCTGCAACATCTTTCATCAACAGATTTTCAGTCCGTCACAAG 261
Db 227 GGGCCCCCGTGAATGTGGCCCTTGCCCTGGAGGTGGCCAGCATCGACCATCTCAGAG 286
QY 262 ACCACATGAGTACCGGGTGAATGTTCTTTTGGGCAACAGTGGGAATGACCCACGCTG 321
Db 287 GCGCAATGAGGTACACCATGACGGTGTCTTCCACAGAGCTGGCGGAGCAGCAGGCTC 346
QY 322 TCTACTCCGAGAATATCTGATGACTCTCTGGACCTCGATCCCTCCATGCTGGACTCTATC 381
Db 347 TCTACT---AACACACAAACGAGACCTTGGGCTGGACAGCGCTTCGTGGACAAGCTG 403
QY 382 TGAAGCCAGACCTCTTTTGTCTAATGAGAAAGGGGCCAACTTCCATGAGGTGACACG 441
Db 404 TGGCTGCCGACACCTTTCATCGTGAAGCCAACTGCGGCTGTTCCACGAGCTGACGGTG 463
QY 442 GACACAAAGTTACTGCGCATCTTCAAGATGGGAATGCTGTACAGCATCAGGCTGACC 501
Db 464 GAGAACAAAGCTCATCCGCTGACGCCCGGAGGTGATCTGTACAGCATCCGAATCACC 523
QY 502 CTCATTTTGTCTGCTGCTGATGACCTCAAGAACTTCCCATGACATCCAGAGCTGCAAG 561
Db 524 TCCACTGTGGCTTGCAGATGAGACCTGGCCAAATTTCCCATGAGCAGGAGGATGCTG 583
QY 562 ATGACAGCTTGAGAGCTCATCTCATCTCTGACGCCCTCTGCCATCTCTGTCACTTTCA 621
Db 584 CTGGACCTGGAGACT-----AC 601
QY 622 GGCTACACCATGAAGACCTCTGTTGATGAGTGGCTGGAAGATGCTCCTGCTGTCCAAGTG 681
Db 602 GGTACTCATCGAGGACATCTGTACTACTGTCGGAGAGCAGGAGCAGCAGCATCCACGG 661
QY 682 GCTGAGGGCTGACTCTGCGCCAGTTTATCTTTCGGGATGAGAAGGATCTAGGCTGTGT 741
Db 662 CTGGCAAGCTGAGCTGGGCGAGTTTCCATCATCCAGCTACCGCTTCCACAGGAGCTG 721
QY 742 ACCAAGCACTACAAACAGGGAAATTCACCTGATCGAGGTAAGTTTTCACCTGGAAACGG 801
Db 722 ATCAACTTCAAGTCCGCTGCGCAGTTTCCACAGGCTCAGCTGACATCCACCTGCGGAG 781
QY 802 CAGATGGGCTACTATCTGATTCAGATGATACATCCCGAGCTTACTCATCTGTCATCTGTC 861
Db 782 AACCGGGGCTGTACATCATCTCAATCTCTATGTCCTGCTGCTGCTGCTGCTGCTGCTGCT 841
QY 862 TGGGTCTCTTCTGGATCAACATGATGCTGCCCTGCGGCTGCTGCTGCTGCTGCTGCTGCT 921
Db 842 TGGGTCTCTTCTGGATCAGCAGGCGGGGTGCCCGGAGGTGCTCTTAGGCATCACC 901

```



```

QY 922 ACCGTGCTCACCATGACCCAGAGCTCTGGCTCCGGGCGCTCTTTGGCTAAGGTGTCC 981
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 902 ACGTCTGACGATGACCAAGCTCATGCTCAGTGGCCCGCTCTCCCTGCCACGGGCATCA 961
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 982 TAGCTGAAGGCAATCGACATCTGGATGGCTGTGTGTCTCTTTGTGTGTGTGTGTG 1041
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 962 GCCATCAAGGCACTGGACGCTCTACTTCTGATCTGTATGTCTTCTGTTTGGCGCCCTG 1021
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1042 CTGGAGTATGC 1052
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1022 GTGGAGTACGC 1032
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 7
US-09-592-891A-13
; Sequence 13, Application US/09592891A
; Patent No. 6329174
; GENERAL INFORMATION:
; APPLICANT: Xiao-Zhou Michelle Wang
; APPLICANT: Xavier Georges Sarda
; APPLICANT: Michael David Tomalski
; APPLICANT: Vincent Paul Mary Wingate
; TITLE OF INVENTION: Heliothis Glutamate Receptor
; FILE REFERENCE: A32815 072667.0118
; CURRENT APPLICATION NUMBER: US/09/592,891A
; CURRENT FILING DATE: 2001-01-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 4621
; TYPE: DNA
; ORGANISM: Heliothis virescens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (764)...(764)
; OTHER INFORMATION: n is a, t, g, or c
US-09-592-891A-13

Query Match 11.4%; Score 187; DB 3; Length 4621;
Best Local Similarity 54.5%; Pred. No. 3.6e-45;
Matches 472; Conservative 0; Mismatches 346; Indels 48; Gaps 3;

QY 208 CCCGTGAACGTGACCTGCAACATCTTCATCAACAGTTTCAGCTCCGTCACCAAGACCA 267
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 330 CCAGCGGTAGTGAGCGTCAATATATTTGTCGGAAGTATATCAAGATCGATGACGTACA 389
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 268 ATGGACTACGGGTGAATGTCTTTGGGCAACAGTGAATGACCCAGCTGTCTCTAC 327
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 390 ATGGAATACCTCCGTACAGTTAAACGTTTCGGGAACAAATGGTTAGATGAACGGCTCAA 449
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 328 CGAGATAATCTCGATGACTCTCTGGACCTCGATCCCTCCATGCTGACTCTATCTGGA 387
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 450 AATAATCTTGGAGTCCGCTCAAAATACCTGACACTGACCCGAAGCCAGAGTCTGGATG 509
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 388 CCAGACTCTCTTTGCTAATGAGAAGGGGCCAACTTCATGAGGTGACCGACGCAAC 447
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 510 CTGTGATCTATTCTTCCACAGGAGGAGGTCAATTCACACATCATCATGCGGCAC 569
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 448 AAGTTACTGGCATCTTCAAGAAATGGAAATGTGTGTAGCATGAGTACAGCTGACCTCAT 507
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 570 GTGTACATCCGGATCTTCCCAACGCAACGCTGTGTACAGCATCCGAATCTCTTTGAG 629
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 508 TTGTCTGCTGATGACCTCAAGAACTTCCCATGAGCATCCAGACGTGCAAGTCAAG 567
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 630 CTCGTGCCCCAATGAACCTCAAGTTGATACCCCTGGTAAGCAGACCTGCTGCTCAG 689
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 568 CTTGAGAGCTCATCCATCTCTGACGCCCTCTGCCATCTCTGCTCACTTTCAGTTGGCTAC 627
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 690 ATGGCTAGTT-----ATGGTTGG 707
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 628 ACCATGAAGACCTCTGTGTTGAGTGGCTGGAGATGCTCTGCTGCTCAAGTGGCTGAG 687
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 708 ACCACAGACGACTTAGTGTCTCTATGGAAGGCGACCC---CGGTGAGGTGGTGAAN 764
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 688 GGGCTGACTCTGCCCGAGTTTATCTTGGCGGATGAGAAGATCTAGGCTGTGTGACCAAG 747
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 765 AACTTACACTGCTCGGTTAC---GCTGGAGAAGTTCTCTACTACTACTGCAACAGT 821
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 748 CACTACAAACACAGGGAATTCACCTGSCATCGAGGTAAAGTTTTCACCTGGAACGCGAGATG 807
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 822 AAGACTAATACCGGTGAATACAGTTGCTGGAAGTAGACTTGTCTCTTCAACGCGAGTTT 881
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 808 GGTACTATCTGATTCAGATGTACATCCCGAGCTCTACTCATGTATCTCTGCTGGTTC 867
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 882 AGTTACTACTGATCCAGATCTACATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 941
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 868 TCCTTCTGATCAACATGATGCTGCCCTGCCCTGCTGGGCTGGGSCATCACACCGTG 927
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 942 TCCTTCTGCTGGACGAGGAGCTGTGCTGCGAGGCTCTCACTAGGAGTGAGACTTTA 1001
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 928 CTCACCATGACCAACCCAGAGCTCTGCTCTCCCGGCTCTTTGCTTAAGGTGCTTACGTG 987
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1002 CTTACAATGGCGACCCAGTCGTGAGGCATCAACGCGTCCCTACCAACGCGTCTTACAG 1061
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 988 AAGGCAATCGACATCTGGATGGCTGTGTGTCTCTTTGCTGCTGCTGCTGCTGCTG 1047
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1062 AAAGCCATTGACGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1121
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1048 TATGCTGCCATAAATTTTGTCTCG 1073
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1122 TTCGCGCTCTCACTATGCGTCTCG 1147
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 8
US-09-969-844-13
; Sequence 13, Application US/09969844
; Patent No. 6780601
; GENERAL INFORMATION:
; APPLICANT: Xiao-Zhou Michelle Wang
; APPLICANT: Xavier Georges Sarda
; APPLICANT: Michael David Tomalski
; APPLICANT: Vincent Paul Mary Wingate
; TITLE OF INVENTION: Heliothis Glutamate Receptor
; FILE REFERENCE: A32815-1 072667.0178
; CURRENT APPLICATION NUMBER: US/09/969,844
; CURRENT FILING DATE: 2001-10-03
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 4621
; TYPE: DNA
; ORGANISM: Heliothis virescens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (764)...(764)
; OTHER INFORMATION: n is a, t, g, or c
US-09-969-844-13

Query Match 11.4%; Score 187; DB 4; Length 4621;
Best Local Similarity 54.5%; Pred. No. 3.6e-45;
Matches 472; Conservative 0; Mismatches 346; Indels 48; Gaps 3;

QY 208 CCCGTGAACGTGACCTGCAACATCTTCATCAACAGTTTCAGCTCCGTCACCAAGACCA 267
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 330 CCAGCGGTAGTGAGCGTCAATATATTTGTCGGAAGTATATCAAGATCGATGACGTACA 389
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 268 ATGGACTACGGGTGAATGTCTTTGGGCAACAGTGAATGACCCAGCTGTCTCTAC 327
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 390 ATGGAATACCTCCGTACAGTTAAACGTTTCGGGAACAAATGGTTAGATGAACGGCTCAA 449
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 328 CGAGATAATCTCGATGACTCTCTGGACCTCGATCCCTCCATGCTGACTCTATCTGGA 387
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 450 AATAATCTTGGAGTCCGCTCAAAATACCTGACACTGACCCGAAGCCAGAGTCTGGATG 509
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 388 CCAGACTCTCTTTGCTAATGAGAAGGGGCCAACTTCATGAGGTGACCGACGCAAC 447
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```
Db 510 CCTGATCTATTCTTCCACGAGAGGAGTCAATTTCCACAAATCATCATGCCGAC 569
Qy 448 AAGTTACTGGCATCTTCAAGAAATGGAATGTCTGTACAGCATCAGGCTGACCCCTCATY 507
Db 570 GTGTACATCCGATCTTCCCAACGCAAGTGTCTGTACAGCATCCGAATCTCTTGACG 629
Qy 508 TTGTCTGCTGTATGATGACCTCAAGAACTTCCCATGAGCATCCAGATGCGATGAG 567
Db 630 CTCCTGCTGCCATGAACCTCAAGTTGTATCCCTGGATAAGCAGACCTGCTGCTCAGG 689
Qy 568 CTTGAGAGCTCATCCATCTCTGCAGCCCTCTGCCATCTCTGTCACTTTTCAGTTGGCTAC 627
Db 690 ATGGCTAGTT-----ATGGTTGG 707
Qy 628 ACCATGAAGACCTCTGTTTGTAGTGGCTGGAAGATGCTCTGCTGTCCAAGTGGCTGAG 687
Db 708 ACCACAGACGACTTAGTGTCTTATGGAAGGAGGCGAC---CGGTGACAGTGGTGAAN 764
Qy 688 GGGCTGACTCTGCCCCAGTTTATCTTGGCGGATGAGAGATCTAGGCTGTTGTACCAAG 747
Db 765 AACTTACACTGCTCGGTTCAC---GCTGGAGAAGTTCTCTCACTGACTACTGCAACAGT 821
Qy 748 CACTCAACACAGGGAATTCACCTGATCGAGGTAAAGTTTCCACCTGGAACGCGAGATG 807
Db 822 AAGACTAATACCGTGATACAGTTGCTGGAAGTACATTTGCTCTTCAACCGAGTTTC 881
Qy 808 GGCTACTATCTGATTCAGATGTACATCCCGAGCTACTCATGCTCATCTCTCTGGTTC 867
Db 882 AGTTACTACTGATCCAGATCTACATTCGCTGTGATGCTGATGCTGCTCTGCTGGTG 941
Qy 868 TCCTTCTGATCAACATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 927
Db 942 TCCTTCTGCTGACGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1001
Qy 928 CTCACCATGACACCCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 987
Db 1002 CTTACAAATGGGACCCAGTGTGCTGAGGATCAACGCTGCCCTACACCGGTGCTTACAG 1061
Qy 988 AAGGCAATCGACATCTGGATGGTGTGTCTGCTCTTGTGTGCTGCTGCTGCTGCTGCTGCT 1047
Db 1062 AAAGCAATGAGCTGTGGAGTGGTGTATGCTCTACATTCGTATTCGGAGGCTACTAGAG 1121
Qy 1048 TATGTCGCAATAAATTTGTTTCTCG 1073
Db 1122 TTGCGCTGCTCAACTATGCTGCTCG 1147
```

## RESULT 9

US-08-417-330A-15  
; Sequence 15, Application US/08417330A  
; Patent No. 5719057

## GENERAL INFORMATION:

; APPLICANT: HADINGHAM, KAREN  
; APPLICANT: LE BOURDELLES, BEATRICE  
; APPLICANT: WHITING, PAUL  
; APPLICANT: WINGROVE, PETER  
; TITLE OF INVENTION: STABLY TRANSFECTED CELL LINE EXPRESSING  
; TITLE OF INVENTION: GABA-A RECEPTOR AND NOVEL CLONED  
; TITLE OF INVENTION: GABA-RECEPTOR SUBUNIT CDNA SEQUENCES  
; NUMBER OF SEQUENCES: 20

## CORRESPONDENCE ADDRESS:

; ADDRESSEE: J. MARK HAND - MERCK & CO., INC.  
; STREET: 126 EAST LINCOLN AVENUE - P.O. BOX 2000  
; CITY: RAHWAY  
; STATE: NJ

## COUNTRY: US

; ZIP: 07065-0900

## COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq For Windows Version 2.0

; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/417,330A  
; FILING DATE: 05-APR-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HAND, MARK  
; REGISTRATION NUMBER: 36,545  
; REFERENCE/DOCKET NUMBER: T11091A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 908-594-3905  
; TELEFAX: 908-594-4720  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2310 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: Coding Sequence  
; LOCATION: 298...1683  
; OTHER INFORMATION:  
; US-08-417-330A-15

Query Match 11.1%; Score 182.2; DB 1; Length 2310;  
Best Local Similarity 54.0%; Pred. No. 6.5e-44;  
Matches 488; Conservative 0; Mismatches 373; Indels 42; Gaps 4;  
Qy 168 TGGATATGATGCCAGATTCGGCCCCAAATTTTAAAGCCCCACCGTGAACGTGACCTGCAA 227  
Db 471 TGGCTACGACAAACAGACTTCGGCCCCGGCTGGGAGAGCGCATCACTCAGGTGAGGACCGA 530  
Qy 228 CATCTTCATCAACAGTTTCAGCTCCGTACCAAGACCAACATGGACTACGGGTGATGT 287  
Db 531 CATCTACGTCAACAGCTTCGGCCCCGGTGTCCGACACGSAATGGAGTACACCATAGACGT 590  
Qy 288 CTTCTTGGCGGCAACAGTGGAAATGACCCACGCTGTCTACCGAGAAATATCTCTGATGACTC 347  
Db 591 GTTTTTCGACAAAGCTGGAAAGATGAAGGCTTCGGTTTAAGGGG---CCCATGACGG 647  
Qy 348 TCTGGACCTCGATCCCTCCATGCTGACCTCTATCTGGAAGCCAGACCTCTTTTGTGCTAA 407  
Db 648 CCTCCCTCTCAACACCTCTTGGCCAGCAAGATCTGGAACCCAGACACAGCTTCTTCCACA 707  
Qy 408 TGAGAAAGGGGCAACTTCCATGAGTGACCGACGACCAACAGTTACTGGGATCTTCAA 467  
Db 708 CGGGAAGAAAGTCCATCGCTCAACACATGACCAACGCGCCCAACAGCTCTCGCGCTGGAGGA 767  
Qy 468 GAATGGGAATGTCTGTACAGATCAGGCTGACCCCTCATTTTGTCTGCTGTATGGACCT 527  
Db 768 CGACGGCACCTGCTCTACACCATGGCTTGACCATCTCTGCAGAGTGCCCATGACGT 827  
Qy 528 CAAGAACTTCCCATGAGCATCCAGACGTGACGATGACGATGAGCTTGGAGCTTCACTACT 587  
Db 828 TGAGGACTTCCCGATGATGCGACGCTTGGCCCTCTGAAATTTGGCAGCT----- 877  
Qy 588 CTGACGCCCTCTGCCATCTCTGTCACTTTCAGTTGGCTACACCATGAAGACCTCTGTGT 647  
Db 878 -----ATGCGTACCCCTAAATCTCTGAAGTCTGTTTACGCTGACCAACAGCGCTCCAC 926  
Qy 648 TGAGTGGCTGGAAGATGCTCTGCTGTCCAAGTGGCTGAGGGGCTGACTCTGCCCCAGTT 707  
Db 927 CAAGTCTGTTGGTGGC-----GGAAGATGGCTCCAGACTGAACCACTA 971  
Qy 708 TATCTTGGCGGATGAGAGGATCTAGGCTGTTGTATCAAGACACTACACACAGGGAAT 767  
Db 972 CCACCTGATGGGGCAGACGG---TGGGCACTGAGACATCAGCACCCAGCACAGCGCAATA 1028  
Qy 768 CACCTGCTCGAGGTAAAGTTTTCACCTGGAAACGGGAGATGGGCTACTATCTGTATTCAGAT 827



```
Db 394 ATGGAATACCTCGTACAAATTAACGTTTCGGGAACAATGGTAGTAAGCGCTCAAAATTC 453
Qy 328 CGAGAATACTCGTATGACTCTCTGSGACCTGGATCCCTCCATGCTGACTCTATCTGGAAG 387
Db 454 AATAAATCTTGGAGGTGCGCTCAAAATACCTGACACTGACTGAAGCCAAAGAGTCTGGATG 513
Qy 388 CCAGACCTCTCTTTCTGTAATGAGAAAGGGGCCAACTTCCATGAGGTGACCGAGCAAC 447
Db 514 CTTGATCTATTCTTCTCCAAAGGAGGAAGGTCAATTCACAAACATCATCATGCGCAAC 573
Qy 448 AAGTTACTGCGCATCTTCAAGAAATGGAATGTCTGTACAGCATCAGGCTGACCCCTCAAT 507
Db 574 GTGTACATCGAATCTTCCCAACGCAAGTGTGTACAGCATCGAATCTCCCTGACG 633
Qy 508 TTGCTCTGCTGATGAGCACTCAAGAAATCTCCCATGAGCATCCAGACGTGCGAGTGCAG 567
Db 634 CTCTCGTGGCCCAATGAACCTCAAGTTGTACCCCTGGATGAAGCAGACCTGCTCGTCAGG 693
Qy 568 CTTGAGAGCTCATCCATACCTCTGCAGCCCTCTGCCATCTCTGTCACCTTCAGTTGGCTAC 627
Db 694 ATGGCTAGTT-----ATGGTTGG 711
Qy 628 ACCATGAAGACCTCGTGTGAGTGGCTGGAAGATGCTCTGCTGTCCAAGTGGCTGAG 687
Db 712 ACCACAGACGACTTAGTGTCTTATGGAAGGAAGGCGACC---CGGTGAGGTGGTGA 768
Qy 688 GGCTGACTCTGCCCGAGTTTATCTTGGCGGATGAGAAGGATCTAGGCTGTTGTACCAAG 747
Db 769 AACTTACACCTGCTCGGTTTCAC---GCTGGAGAAGTTCCTCACTGACTACTGCAACAGT 825
Qy 748 CACTACAACAGCGGAATTCACCTGATGAGGTAAAGTTTCACTGGAACCGGAGATG 807
Db 826 AAGACTAATACCGGTGAATACAGTTGCTGAAGGTAGACCTGCTCTTCAAAACGCGAGTTC 885
Qy 808 GGCTACTATCTGATTCAGATGTATCTCCAGCCTACTCATCTCATCTGCTCTGCTGGTTC 867
Db 886 AGTTACTACTGATTCAGATCTACATTCCTGCTGATGCTGCTGATGCTGCTGCTGCTGGTG 945
Qy 868 TCCTTCTGATCAACATGATGATGCTGCCCTGCTGCGGCTGTTGGGCTGAGGCTACCAACCGTG 927
Db 946 TCCTTCTGCTGGAACGAGGAGTGTGCTGCGAGGCTCTCACTAGGAGTGAAGTCTTA 1005
Qy 928 CTCACATGACACCGAGAGCTGCTGCTGCCGGGCTCTTTGCTGAAGGTGCTCAAGTG 987
Db 1006 CTTACAATGGCGACCCAGTCTGAGGATCAACGCGTCCCTACCAACCGGTGCTCTACAG 1065
Qy 988 AAGGCAATCGACATCTGGATGGCTGTGTCTGCTCTTTGTGTTGCTGCTGCTGCTGGAG 1047
Db 1066 AAGCCATTGATGCTGGAGCTGGGTATGTCTCAATTCGATTCGGAGCGCTACTAGAG 1125
Qy 1048 TATGCTGCCATMAATTTTGTCTCG 1073
Db 1126 TTGCGCTGCTCAACTATGGCTCG 1151
```

## RESULT 12

```
US-08-435-933-5
; Sequence 5, Application US/08435933
; Patent No. 5693492
; GENERAL INFORMATION:
; APPLICANT: Cully, Doris F.
; APPLICANT: Arena, Joseph P.
; APPLICANT: Pares, Philip S.
; APPLICANT: Liu, Ken K.
; TITLE OF INVENTION: DNA ENCODING GLUTAMATE GATED CHLORIDE
; TITLE OF INVENTION: CHANNELS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John W. Wallen III
; STREET: 126 East Lincoln Avenue
; CITY: Rahway
; STATE: New Jersey
```

```
; COUNTRY: US
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/435,933
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Wallen, III John W.
; REGISTRATION NUMBER: 35,403
; REFERENCE/DOCKET NUMBER: 19264
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 594-3905
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3958 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-435-933-5
```

```
Query Match 10.9%; Score 179; DB 1; Length 3958;
Best Local Similarity 53.7%; Pred. No. 8.4e-43;
Matches 473; Conservative 0; Mismatches 360; Indels 48; Gaps 3;

Qy 193 AATTTTAAAGGCCACCCGCTGAACGTGACCTGCAACATCTTCATCAACAGTTTCAGCTCC 252
Db 570 AATGGCACAGATGGTCCGCCATAGTCAGATCAATCTATTCTGTCGCAATATTATGACG 629
Qy 253 GTCACCAACAGCACATGACTACCGGTGAATGCTCTTCTGCGGCAACAGTGAATGAC 312
Db 630 ATTAGTGATTTAAATGAGGTACAGTGTGAGTTAACTTCCGTGAACAGTGGACGGAT 689
Qy 313 CCACGCTGCTCTTACCGAGAATATCTGATGACTCTCTGGACCTCGATCCCTCATGCTG 372
Db 690 GAACGCTCAAGTTCGACGATATCCAGGCTCGCTAAAGTATCTGACCTGACGGAGCG 749
Qy 373 GACTCTATCTGAAGCCAGACCTCTTCTTTGCTAATGAGAAAGGGGCAACTTTCATGAG 432
Db 750 AACCGGTGTGATGCGCGATCTTTTCTCTCGAAGAGAGGAGGACACTTCCCAAC 809
Qy 433 GTGACACCGGACCAACAAGTTACTGCGCATCTTCAAGAATGGGAATGCTGTGACAGATC 492
Db 810 ATCATCATCCCAATGTGTATATTCGCATCTTCCCAACCGGATCTGTGCTATATAGTATA 869
Qy 493 AGGCTGACCCCTCATTTTGTCTGCTGCTGATGGACCTCAAGAACTTCCCATGGACATCCAG 552
Db 870 CGTATCTCGCTGACATTTGGCTGCGCAATGAACCTAAAGCTGTATCGCTGGATAGACAG 929
Qy 553 ACGTGACAGATGAGCTTGAAGAGTCAATCCATCTGACGCCCTCTGCCATCTCTGTCTCA 612
Db 930 ATCTGCTCACTACGATGCGGCTGCT-----954
Qy 613 CTTTCAGTTGGCTACACCATGAAGACCTCTGTTTGTGCTGGCTGGAAGATGCTCTGCT 672
Db 955 -----ATGGCTGGACCAACCAAGCTTGGTCTTCTCTGGAAGAGGCGGATC---CC 1004
Qy 673 GTCCAAAGTGGCTGAGGGGCTGACTCTGCGCCCAAGTTTATCTTGGGATGAGAAGGATCTA 732
Db 1005 GTACAGTGTGAAGAACTTACACCTACCTGCTTCACTGGAGAGTTTCTGACT--- 1061
Qy 733 GGCTGTTGTACCAAGCACTACCAACAGAGAAATTCACCTGCAATCAGGTAAAGTTTCAC 792
Db 1062 GATTACTGTAAACAGTAAACCAACACCGGTGAATACAGTTGCTCAAGTCTGATCTACTA 1121
Qy 793 CTGGAAACGCGAGATGGGCTACTATCTGATTCAGATGTATATCCCAAGCTACTCATCTGC 852
```

Db 1122 TTCAGCGGAGAAATTCATATATTAATAAATAATTAATACATGCTGTATGTGGTC 1181  
 QY 853 ATCTGTCTGGGTCTCTTCTGGATCAACATGATGCTGCCCTCCCGTGGGCGTG 912  
 Db 1182 ATTGTATCATGGTATCATCTTGGCTGGATCAAGAGAGATACCGGCGGAGTGTACAG 1241  
 QY 913 GGATCACCACCGTGTCTACCATGATGACCCAGAGCTCTGGCTCCCGGCGCTCTTTGGCT 972  
 Db 1242 GGTGTCAACACCGTGTCTGACCATGCGCCAGAGCTGGGCATAAAGCCCTCCCTGCG 1301  
 QY 973 AAGGTGTCTAGTGAAGCAATCGACATCTGGATGGCTGTGTCTGTCTTTGTGTTC 1032  
 Db 1302 CCCGTTTCTATACGAAGGCCATCGATGTGTGACAGGGGTGTGTCTGACGTTCTGTTC 1361  
 QY 1033 GCTGCTTCTGAGTATGCTGCCATAAATTTTGTTCG 1073  
 Db 1362 GGGGCCCTGCTGAGTTCGCCCTGGTGAATGATCCCG 1402

RESULT 13

PCT-US96-06035-5  
 ; Sequence 5: Application PC/TUS9606035  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cully, Doris F.  
 ; APPLICANT: Arena, Joseph P.  
 ; APPLICANT: Pares, Philip S.  
 ; APPLICANT: Liu, Ken K.  
 ; TITLE OF INVENTION: DNA ENCODING GLUTAMATE GATED CHLORIDE  
 ; TITLE OF INVENTION: CHANNELS  
 ; NUMBER OF SEQUENCES: 6  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Jody M. Giesser  
 ; STREET: 126 East Lincoln Avenue - P.O. BOX 2000-0907  
 ; CITY: Rahway  
 ; STATE: New Jersey  
 ; COUNTRY: US  
 ; ZIP: 07065-0907  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent in Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US96/06035  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Giesser, Jody M.  
 ; REGISTRATION NUMBER: 32,838  
 ; REFERENCE/DOCKET NUMBER: 19264 PCT  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (908) 594-3046  
 ; TELEFAX: (908) 594-4720  
 ; INFORMATION FOR SEQ ID NO: 5:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 3958 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; PCT-US96-06035-5

Query Match 10.9%; Score 179; DB 5; Length 3958;  
 Best Local Similarity 53.7%; Pred. No. 8.4e-43;  
 Matches 473; Conservative 0; Mismatches 360; Indels 48; Gaps 3;  
 QY 193 AATTTTAAAGGCCACCCGTGAACGTGACATCTTCAACATCTTCAACAGTTTCAAGCTCC 252  
 Db 570 AATGGCAGAGTGGTCCCGCCATGATCAATCTATTCTGTCGAGTATATGACG 629  
 QY 253 GTCACCAAGCCCAATGACACTACCGGGTGAATGTCTTCTGGGCAACAGTGGATGAC 312  
 Db 630 ATTAGTGATTAATAATGGAGTACAGTGTGCGATTAACTTCCGTGAACAGTGGACGGAT 689

QY 313 CCACGCGTGTCTTACCGAGAAATATCTGTATGACTCTCTGGACCTCGATCCCTCCATGCTG 372  
 Db 690 GAAAGCGCTCAAGTTCGACGATATCCAGGTCCCTAAAGTATCTGACCTCGAGGAGCG 749  
 QY 373 GACTTATCTGAAGCCAGACCTCTTCTTGTCTAATAGAAAAGGGGCCAACTTCCATGAG 432  
 Db 750 AACCAGGTGTGATGCCGATCTTTTCTCTGAAAGAGAGGAGGACACTTCCACAAC 809  
 QY 433 GTGACCCAGGACAAAGTCTTACTGCGCATCTTCAAGAAATGGAATGTGCTGTACAGCATC 492  
 Db 810 ATCATCATGCCCAATGTGTATATTCGCATCTTCCCAACGGATCTGTGTATATAGTATA 869  
 QY 493 AGGCTGACCTCATTTTGTCTGCTGATGAGACCTCAAGAACTTCCCCATGACATFCCAG 552  
 Db 870 CGTATCTCGTCACATTTGGCTGCCCAATGAACCTAAAGCTGTATCCGCTGGATAGACAG 929  
 QY 553 ACGTGACGATGACGCTTGAGAGCTCATCATCTCTGAGCCCTTGCCATCTCTGTCTCA 612  
 Db 930 ATCTGCTCACTACGGATGGCCAGCT----- 954  
 QY 613 CTTTCAGTTGGCTACACCATGAAGACCTCGTGTGTGAGTGGCTGGAAGATGCTCTGTCT 672  
 Db 955 -----ATGGCTGACCAACCAACGACTTGGTCTTCTGTGGAAGAGGCGGATC---CC 1004  
 QY 673 GTCCAAGTGGCTGAGGGGCTGACTCTGCCCAAGTTTATCTTCCGGGATGAGAAGGATCTA 732  
 Db 1005 GTACAGTGGTAAAGAACTTACACCTACCTCGCTTCACTGAGAGAAGTTTCTGACT--- 1061  
 QY 733 GGCTGTGTGACCAAGCACTTACAAACAGAGAAATTCACCTGTCATGAGTAAAGTTTTCAC 792  
 Db 1062 GATTACTGTAAACAGTAAACCAACACCGGTGAATACAGTTGCTCAAAGTTCGATCTACTA 1121  
 QY 793 CTGGAACGGCAGATGGGCTACTATCTGATTCAGATGATCATCCAGCCTACTCATCGTC 852  
 Db 1122 TTCAGCGGAGAAATTCATATTAATAAATAATTAATACATGCTGTATGTGGTC 1181  
 QY 853 ATCTGTCTGGGTCTCTTCTGAGTCAACATGAGTGTGCTGCCCTCCCGTGTGGGCGTG 912  
 Db 1182 ATTGTATCATGGTATCATCTTGGCTGGATCAAGAGCAGTACCGGCGGAGTGTCACTG 1241  
 QY 913 GCATCACCACCGTGTCTACCAATGACCCAGAGCTGTGGCTCCCGGCGCTCTTTGGCT 972  
 Db 1242 GGTGTCAACACCGTGTGACCATGCGCCAGACCCAGAGCTGCGGCATAAACGCTCCCTGCG 1301  
 QY 973 AAGTGTCTAGTGAAGCAATGACATCTCGATGGCTGTGTCTGTCTCTTTGTGTTC 1032  
 Db 1302 CCCGTTTCTATACGAAGGCCATCGATGTGTGACAGGGGTGTGTCTGACGTTCTGTTC 1361  
 QY 1033 GCTGCTTGTGAGTATGCTGCCATAAATTTTGTTCG 1073  
 Db 1362 GGGGCCCTGCTGAGTTCGCCCTGGTGAATGATCCCG 1402

RESULT 14

US-09-949-016-16344  
 ; Sequence 16344, Application US/09949016  
 ; Patent No. 6812339  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig et al.  
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
 ; FILE REFERENCE: CL001307  
 ; CURRENT APPLICATION NUMBER: US/09/949,016  
 ; PRIOR FILING DATE: 2000-04-14  
 ; PRIOR APPLICATION NUMBER: 60/241,755  
 ; PRIOR FILING DATE: 2000-10-20  
 ; PRIOR APPLICATION NUMBER: 60/237,768  
 ; PRIOR FILING DATE: 2000-10-03  
 ; PRIOR APPLICATION NUMBER: 60/231,498  
 ; PRIOR FILING DATE: 2000-09-08  
 ; NUMBER OF SEQ ID NOS: 207012  
 ; SOFTWARE: FastSeq for Windows Version 4.0

```

; SEQ ID NO 16344
; LENGTH: 73519
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: missing
; LOCATION: (1)
; OTHER INFORMATION:
US-09-949-016-16344

```

	Query Match	10.8%;	Score 176.8;	DB 4;	Length 73519;
	Best Local Similarity	87.7%;	Pred. No. 2.4e-41;		
	Matches 193;	Conservative 0;	Mismatches 27;	Indels 0;	Gaps 0;
Qy	758	CAGGGAAATTCACCTGCATCGAGGTAAGTTTCACCTCGGAACGGCAGATGGGCTACTATC	817		
Db	42818	CAGGTAAATTCACCTGCATTCGAGCCCGGTTCCACCTCGAGCGGCAGATGGGTACTACC	42877		
Qy	818	TGATTCAGATGATACATCCCGACGCTACTCATCGTATATCTGCTCGGGTCTCCTTCTCGGA	877		
Db	42878	TGATTCAGATGATATATTCGCAGCGCTGCTCATTTGTATCTCTCATGGATCTCCTTCTCGGA	42937		
Qy	878	TCAACATGATGATGTCGCCCTGCCGTCGGCTGGGCATACCAACCGTGTCCACCATGA	937		
Db	42938	TCAACATGATGTCGCACCTGCTGTGTGGGCTAGGCATACCACTGTGCTCACCATGA	42997		
Qy	938	CCACCCAGAGCTCTGGCTCCCGGGCCTCTTTGCCTAAGGT	977		
Db	42998	CCACCCAGAGCTCCGGCTCTCGAGCATCTCTGCCAAGGT	43037		

## RESULT 15

```

US-09-949-016-11769
; Sequence 11769, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CU001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11769
; LENGTH: 105919
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(105919)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-11769

```

	Query Match	10.8%	Score 176.8;	DB 4;	Length 105919;
	Best Local Similarity	87.7%;	Pred. No. 3e-41;		
	Matches 193;	Conservative 0;	Mismatches 27;	Indels 0;	Gaps 0;
Qy	758	CAGGGAAATTCACCTGCATCGAGGTAAAGTTTCACCTGGAAACGGCAGATGGGCTACTATC	817		
Db	75219	CAGGTAATTCACCTGCATGTAGGCCCGGTTCCACCTGGAGCGGCAGATGGGTATTACTACC	75278		
Qy	818	TGATTTCAGATGTACATCCCGCAGCCTACTCATCGTCACTCGTCTGGGGTCTCCTTCTGGA	877		
Db	75279	TGATTTCAGATGTATATTTCCGAGCCTGCTCATTTGTCAATCTCTCATGTGATCTCCTTCTGGA	75338		
Qy	878	TCAACATGGATGTCGCCCTCTGCCCGGTGTGGGGCATACCCACCGTGCTCACCATGGA	937		

**This Page Blank (uspto)**



Qy







		Query Match	49.3%; Score 807.8; DB 14; Length 993;
		Best Local Similarity	90.0%; Pred. No.6-249;
		Matches 931; Conservative	0; Mismatches 2; Indels 102; Gaps 2;
Qy	1	ATGACAACCTGTGGTTCCTTGCAAAACCCTCTCCTCCCTTTCTCTCGAGCCCTGCCAAGGCAG	60
Dd	1	ATGACAACCTGTGGTTCCTTGCAAAACCCTCTCCTCCCTTTCTCTCGAGCCCTGCCAAGGCAG	60
Qy	61	GTCCTCTCAGGGTGGCCTTGGCAAAGAAGGAAGTAATACTGGAACCAAGGGTCCCAG	120
Dd	61	GTCCTCTCAGGGTGGCCTTGGCAAAGAAGGAAGTAATACTGGAACCAAGGGTCCCAG	120
Qy	121	CCCATGTCCCCCTCTGATTCTTAGAACAAACTTATGGGGCGAAACATCTGGATATGATGCC	180
Dd	121	CCCATGTCCCCCTCTGATTCTTAGAACAAACTTATGGGGCGAAACATCTGGATATGATGCC	180
Qy	181	AGGATTCGGGCCCAATTTTTAAAGGCCCACCGGTGAACGTGACCTGCAACATCTTCATCAAC	240
Dd	181	AGGATTCGGGCCCAATTTTTAAAGGCCCACCGGTGAACGTGACCTGCAACATCTTCATCAAC	240
Qy	241	AGTTTCAGCTCCGTCACCAAGACCAACAT-----	269
Dd	241	AGTTTCAGCTCCATCACCAAGACCAACATGSGCTTGTGGGCCCTGGGAATGGCAATGTT	300
Qy	270	-----GGACTACCGGGTGAATGTCTTCTTGGCGCA	300
Dd	301	TCTGAAGGGGCCAPATCTGCAACCTCCAGGACTACCGGGTGAATGTCTTCTTGGCGCA	360
Qy	301	CAGTGGAAATGACCAACGCTGTCTACCGAGAATATCTGNATGACTCTCTGGACCTCGAT	360
Dd	361	CAGTGGAAATGACCAACGCTGTCTACCGAGAATATCTGNATGACTCTCTGGACCTCGAT	420
Qy	361	CCCTCCAATGCTGGACTATCTGGAAGCCAGACCTCTTCTTGTGTAATGAAAAGGGGCC	420
Dd	421	CCCTCCAATGCTGGACTATCTGGAAGCCAGACCTCTTCTTGTGTAATGAAAAGGGGCC	480
Qy	421	AAC TTCCA TAGAGTGA CCA CGGA CA A AG TTA CTGG GCATCTTTCA GA ATGG AAATGTG	480
Dd	481	AAC TTCCA TAGAGTGA CCA CGGA CA A AG TTA CTGG GCATCTTTCA GA ATGG AAATGTG	540
Qy	481	CTGTACAGCATCAGGCTGACCCCTCATTTTGTCCCTGCTGATGGACCTCMAGAACTTCCCC	540
Dd	541	CTGTACAGCATCAGGCTGACCCCTCATTTTGTCCCTGCTGATGGACCTCMAGAACTTCCCC	600
Qy	541	ATGGACATCCAGACGTCGACGATGCAGCTTGAGAGCTCATCCATCTCTGCAGCCCTCTG	600
Dd	601	ATGGACATCCAGACGTCGACGATGCAGCTTGAGAGCT-----	637
Qy	601	CCATCTCTGTC ACTTT CAGTTGGTCA CCACATGA AAG AC CTC GTTGTGAGTGGCTGAA	660
Dd	638	-----TTGGGCTAC ACCATGA AAG ACC TC GTTGTGAGTGGCTGAA	678
Qy	661	GATGCTCTGCTGTCCAAGTGGCTGAGGGGCTGACTCTGCCCGCAGTTTATCTTCCGGGAT	720
Dd	679	GATGCTCTGCTGTCCAAGTGGCTGAGGGGCTGACTCTGCCCGCAGTTTATCTTCCGGGAT	738
Qy	721	GAGAAGGATCTAGGCTGTGTGTACCAAGCACTFACAACACAGGGAAATTCACCTGCATCGAG	780
Dd	739	GAGAAGGATCTAGGCTGTGTGTACCAAGCACTFACAACACAGGGAAATTCACCTGCATCGAG	798
Qy	781	GTAAAGTTTCA CCTGG AACGGCAGATGGCTACTATCTGATTCAGATGTATACATCCCCAGC	840
Dd	799	GTAAAGTTTCA CCTGG AACGGCAGATGGCTACTATCTGATTCAGATGTATACATCCCCAGC	858
Qy	841	CTACTCATGTCATCTCTGTCCTGGGTCTCCTTCTGGATCAACATGGATGCTGCCCTGCC	900
Dd	859	CTACTCATGTCATCTCTGTCCTGGGTCTCCTTCTGGATCAACATGGATGCTGCCCTGCC	918
Qy	901	CGTGTGGGCTGGGSCATCACACCGTGTCAACATGACCACCCAGAGCTCTGGCTCCCGG	960
Dd	919	CGTGTGGGCTGGGSCATCACACCGTGTCAACATGACCACCCAGAGCTCTGGCTCCCGG	978
Qy	961	GCCTCTTTGGCTAAG	975

```

Db          979 G0CTCTTGCCTAAG 993
|||||
RESULT 5
US-10-757-262-9
; Sequence 9, Application US/10757262-9
; Publication No. US20040197825A1
; GENERAL INFORMATION:
; APPLICANT: Karicheti, Venkateswarl
; APPLICANT: Siles-Santiago, Inmacu
; APPLICANT: Eliasof, Scott D.
; TITLE OF INVENTION: METHODS AND COM
; TITLE OF INVENTION: UROLOGICAL DI
; TITLE OF INVENTION: 1405, 636, 44
; TITLE OF INVENTION: 33751, 52872,
; TITLE OF INVENTION: 53010, 16852,
; TITLE OF INVENTION: 18547, 115, 5
; TITLE OF INVENTION: 32720, 4809,
; TITLE OF INVENTION: 2158, 8463, 1
; TITLE OF INVENTION: 2543, 9626, 1
; TITLE OF INVENTION: 55053
; FILE REFERENCE: MPI03-007P1RNMOMIM
; CURRENT APPLICATION NUMBER: US/10/
; CURRENT FILING DATE: 2004-01-14
; PRIOR APPLICATION NUMBER: US 60/44
; PRIOR FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: US 60/44
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: US 60/45
; PRIOR FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: US 60/46
; PRIOR FILING DATE: 2003-03-08
; PRIOR APPLICATION NUMBER: US 60/47
; PRIOR FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: US 60/47
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: US 60/48
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/49
; PRIOR FILING DATE: 2003-07-30
; PRIOR APPLICATION NUMBER: US 60/49
; PRIOR FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US 60/50
; PRIOR FILING DATE: 2003-09-26
; NUMBER OF SEQ ID NOS: 136
; SOFTWARE: FastSeq for Windows Vers
; SEQ ID NO 9
; LENGTH: 1715
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (297)...(1646)
US-10-757-262-9

```

	Query Match	40.0%	Score 656;	DB 18;	Length 1715;
	Best Local Similarity	71.9%;	Pred. NO. 7.7e-200;		
	Matches 95;	Conservative 0;	Mismatches 315;	Indels 57;	Gaps 5;
Qy	89	AGGAAGTCAAATCTGGAA	CCAAAGGGGTCCAGGCCCATGTCCCCCTCTGATTTTCCTAGACA	148	
Db	367	AGGAGGCTGAAGCTGCTCGCTCCGC	AACCAAGCCTATGACCCCTCGGATTTCTGGATA	426	
Qy	149	AACATTATGGGGCGAA	CATCTTGGATATGATGCCAGGATTCGGCCCAATTTTAAAGGGCCAC	208	
Db	427	AGCTAATGGGGGAGAA	CTCCGGATATGATGCCAGGATCAGGCCCAATTTTAAAGGTCCCC	486	
Qy	209	CCGTGAACGTGAC	TCGGAACATCTTCATCAAGTTTCAGTCCGTCCAGACCAACACAA	268	
Db	487	CAGTGAACGTGAGCTGC	AACATTTTCATCAACAGCTTTGGTTCATTCGTGTGAGACAACCA	546	
Qy	269	TGGACTACCGGGGTGAATGTCTTCTTGGCGGCAACAGTGGAA	TGATGACCAACCGCTGTCTTACC	328	

547 TGGACTATAGGTCACATCTTCTCGGCAGCAATGGAACGACCCCGCTGGCCTATA 606  
 329 GAGAAATATCTGATGACTCTCTGGACTCGATCCCTCCATCTGCTGGACTCTATCTGAAGC 388  
 607 ATGAATACCTGACGACTCTCTGGACTCGACCTCCATCTGCTGGACTCTATCTGGAAC 666  
 389 CAGACCTCTTCTTTGCTAATGAGAAAGGGCCAACTTCCATGAGGTGACACGGAACA 448  
 667 CTGACCTGTTCTTTGCAACGAGAGAGGGGCCACTTCCATGAGATCACACAGACA 726  
 449 AGTTACTGGGCATCTTCAAGATGGAATGTGCTGTACAGCATCAGGCTGACCTCATTT 508  
 727 AATTGTAAGGATCTCCCGAATGGAATGTCTCTACAGCATCAGATCAACCTGCAC 786  
 509 TGTCTGCTGATGAGACCTCAAGAACTCCCATGACATCCAGCTGCAAGATCAGC 568  
 787 TGGCCTGCCCCATGGACTTGAAGATTTCCCATGATGTCAGACATGATCATCAAC 846  
 569 TTGAGAGCTCATCCATCTCTGAGCCCTCTGCCATCTCTGTCACTTTCAGTTGGCTACA 628  
 847 TGAAGAGCT-----TTGGATATA 864  
 629 CAGTAAAGACCTGTTTGTAGTGGCTGGAAGATGCTCTGCTGCCAAGTGGCTGAGG 888  
 865 CGATGAATGACCTCATCTTTGAGTGGCAGGAACAGG---GAGCGGTGCAGGTAGCAGATG 921  
 689 GGTGACTCTGCCCACTTATCTCTGGGATGAGAGGATCTAGGCTGTGTACCAAGC 748  
 922 GACTAATCTGCCCACTTATCTTGAAGAGAGGAGGACTTGAATGACTGACCAACG 981  
 749 ACTACAAACAGGAAATCACCTGATCGAGTAAAGTTTCACTTGGAAACGCGCATGG 808  
 982 ACTACAAACAGGTAATTCACCTGATTCAGGCCCGGTTCACCTGGAGGCGCATGG 1041  
 809 GCTACTATCTGATTCAGATGATATCCCGAGCTCTCATGTCATCTCTGCTGCTGCT 868  
 1042 GTTACTACCTGATTCAGATGATATTCAGGCTGCTCATTTGTCATCTCTCATGATCT 1101  
 869 CTTCTGGATCAACATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 928  
 1102 CTTCTGGATCAACATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1161  
 929 TCACCATGACCCAGAGCTCTGGCTCCGGGCTCTTTTGCCTAAGGTGTCTTACGTGA 988  
 1162 TCACCATGACCCAGAGCTCCGGCTCTCGACATCTCTGCCAAGGTGCTCTATGTA 1221  
 989 AGCAATTCAGATCTGGATGGCTGTGCTGCTCTTTGTTGCTGCTGCTGCTGCTGCTGCT 1048  
 1222 AAGCCATTCAGATTCGATGGAGTTTGCCTGCTCTTTGTTGCTCTCAGCCCTATTAGAAT 1281  
 1049 ATGCTGCCATTAATTTGTTCTGTCAGATTAAGATTTTCACTGCTGCTGCTGCTGCTGCT 1108  
 1282 ATGCTGCCATTAATTTGTTCTGTCAGATTAAGATTTTCACTGCTGCTGCTGCTGCTGCT 1341  
 1109 AGAGGGCCCAACGCTTGGAGGAAGATATCATCAAGAAAGTCTGTTTCTATTTCCGTGGCT 1168  
 1342 GGAGATCACAGGAGGATGAAGCT-----GGAGAGCCGCTTTAACTTCTGCTGCT 1395  
 1169 ATGGCTTGGGCC---ACTGCTGCGAGCAAGATGAGGTGCTCAATGGAAGGTTTC---TG 1222  
 1396 ATGGGATGGGCCCGAGCTGCTTACAGGCCAAGGATGCGATCTCAGTCAAGGGGCGCAACA 1455  
 1223 GCATTTATAGTCCCAACCTCCAGCCCTCTTCTTAAGGAAGAGAGAAACACCGCGAAGC 1282  
 1456 ACAGTAACACCAACACCCCTCTCTGACCATCTTAAGTCCCGAGAGAGATGCGAAGC 1515  
 1283 TCTACGTGGACTGAGCAAGAGAAATGACACCATCTCCCGGGCTGTCTTCCCTTTTCACTT 1342  
 1516 TCTTCATCAGAGGGCCAGAGATCGACAAATATCCCGCATTTGGCTTCCCGCATGGCCT 1575  
 1343 TCCTCATCTTCAATATCTTACTGGGTGCTTATTAAGTGTCTATGGTCAGAGATATCC 1402

1576 TCTCATTTTCAACATGTTTCTACTGATCATCTACAAGATTGTCCTAGAGAGACGTCC 1635  
 1403 ACCA 1406  
 1636 ACNA 1639  
 RESULT 6  
 US-10-205-194-20  
 ; Sequence 20, Application US/10205194  
 ; Publication No. US20030134301A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Warner-Lambert Company  
 ; APPLICANT: Lee, Kevin  
 ; APPLICANT: Dixon, Alister  
 ; APPLICANT: Brooksbank, Robert  
 ; APPLICANT: Pinnock, Robert  
 ; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain  
 ; FILE REFERENCE: WL-A-018201  
 ; CURRENT APPLICATION NUMBER: US/10/205,194  
 ; CURRENT FILING DATE: 5200-07-24  
 ; PRIOR APPLICATION NUMBER: GB 0118354.0  
 ; PRIOR FILING DATE: 2001-07-27  
 ; NUMBER OF SEQ ID NOS: 177  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 20  
 ; LENGTH: 3865  
 ; TYPE: DNA  
 ; ORGANISM: Rattus norvegicus  
 ; FEATURE:  
 ; OTHER INFORMATION: Neonatal glycine receptor  
 US-10-205-194-20  
 Query Match 39.9%; Score 654.8; DB 15; Length 3865;  
 Best Local Similarity 70.8%; Pred. No. 2.9e-199;  
 Matches 938; Conservative 0; Mismatches 332; Indels 54; Gaps 3;  
 84 AAAAGAGGAGGTCAAATCTGGAACCAAGGGGTCCAGAGCCCATGTCCCTCTGATTTCTCT 143  
 1256 AGACATGATCCAGTCTGGAACCATCCCTCGCAGACCCCTGCTCTCTCAGATTTCTT 1315  
 144 AGACAACTTATGGGGGGAACATCTGGATATGATGCGAGGATTCGSCCAATTTTAAAG 203  
 1316 GGATAACTTAATGGGAAGGACATCAGGATATGATGCAAGAAATCAGGCCAAATTTTAAAG 1375  
 204 CCCACCGTGAAGTGCACCTGCAACATCTTCAACAGTTTTCAGCTTCCTGTCACCAAGAC 263  
 1376 TCTCCAGTAAACGTTACTTGCAATATTTTATCAACAGTTTGGATCGGTACAGAAC 1435  
 264 CACAATGGAATACCGGGTGAATGTTCTTCTGGGCAACAGTGAATGACCCAGCCTGTC 323  
 1436 CACCATGGAATACCGAGTGAACATTTTCTGAGACAGCAGTGAACGATTCAGCGCTGGC 1495  
 324 CTACCGAGAAATATCTGATGACTCTCTGGAACCTGATTCCTGATGCTGCTGATCTATCTG 383  
 1496 ATACAGTGAATACCGAGATGATTCCTGATTTGGATCCATGCTGATGCTGATTTG 1555  
 384 GAAGCAGACCTCTCTTTGCTAATGAGAAAGGGGCAACTTCCATGAGTGAACACGGA 443  
 1556 GAACCAAGATTTGTTCTTTGCGCAATGAGAAAGAGCCAAATTTCCATGATGTCACCACTGA 1615  
 444 CAACAGTTTACTGCGCATCTTCAAGAAATGGGAATGTGCTGTACAGCATCAGGCTGACCT 503  
 1616 TAACAGTTGCTGCGGATTTCCAAATAATGCAAAAGTGTCTACAGTATTAGACTCACCTT 1675  
 504 CATTTTGTCTGCTGATGAGCTCAAGAACTTCCCATGGAATCCAGACGTGACGAT 563  
 1676 GACTTTATCTGTCCTGAGACCTGGAAGAACTTTTCAATGGATGTCCAGACCTGTCAAT 1735  
 564 GCAGCTTGAAGAGCTATCCATCTCTGAGCCCTCTGCAATCTCTGCTCACTTTTCACTTGG 623  
 1736 GCAGCTGAGAGTT-----TTGG 1753



QY 624 CTACACCATGAAGACCTCGTGTGAGTGGCTGGAAGATGCTCCTGCTGCTCAAGTGGC 683  
 Db 1754 GTACACCATGAAGACCTGATATTTGAGTGGTAAAGTATGGTGC---CAGTACAAGTGGC 1810  
 QY 684 TGAGGGGCTGACTGTGCCAGATTTATCTTTCGGGATGAGAAGATCTAGGCTGTGTAC 743  
 Db 1811 TGAAGGACTCACCTGCTCAGTTTATTTTGAAGAAGAGAACTGGCTATTGCAC 1870  
 QY 744 CAAGCACTACACACAGGGAAATTCACCTGCATCGAGGTAAAGTTTCACTCGAAGCGCA 803  
 Db 1871 AAAGCAATTAACACTGGCAAGTTTACCTGATTTAGGTCAAGTTTCACTCGAAGCGCA 1930  
 QY 804 GATGGGCTACTATCTGATTCAGATGTACATCCAGCTACTCATCTGCTATCTGCTCCG 863  
 Db 1931 GATGGGCTATTATTTGATTCAGATGTATATCCAGCTCTCTGATAGTCAATTTGCTCG 1990  
 QY 864 GGTCTCCTCTCGATCAACATGATGCTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 923  
 Db 1991 GGTCTCCTCTTGGATAAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2050  
 QY 924 CGTCTCCTGATGACATCGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 983  
 Db 2051 CGTCTCCTGATGACTACCCAGAGTTCAAGTTTCCAGGGCATCTCTGCCAAAGTCTCATA 2110  
 QY 984 CGTGAAGCAATCGACATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1043  
 Db 2111 TGTGAAGCAATGACATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2170  
 QY 1044 GGATGATGCTGCCATAAATTTGTTTCTCGTACGATAAAGAAATTCATAGCACTTCGAA 1103  
 Db 2171 GGAATATGACAGTGNACTTTGTTCTCCAGGCAACACAGGAGTTCTTCTGCTCCGGAG 2230  
 QY 1104 AAGCGAGCGCCAAAGCTTGGAGGAAGATATCATCAAGAAAGTGGTTTCTATTTCGG 1163  
 Db 2231 ACACAGAGAGGAGCAATTAAGCAAGAGATGTTACTCGTGAAGCGGTTTAACTTCAG 2290  
 QY 1164 TGCTATGCTTGGGCACTGCTGAGGCAAGATGAGTGGAGTCCATGGAAGTTCTGG 1223  
 Db 2291 CGGTTATGGGATGGTCACTGCTCCAGAGTGAAGATGCTACAGCTGTCAAGGCTAC--- 2347  
 QY 1224 CATTTATAGTCCCACTCCAGCCCTCTTCAAGGAAGGAGAAACACCGGAAACT 1283  
 Db 2348 -----ACCTGCCAACCCACTTCGGCAACCCCAAGATGAGATGCTATCAAGAGAA 2401  
 QY 1284 CTACGTGAGTACGCAAGAGAAATGACACCATCTCCCGGCTGTCTTCCCTTTCACTTT 1343  
 Db 2402 GTTTGTGATCGGCAAGAAAGATCGACACCATCTCGAGTGGCTTCCCACTGGCCTT 2461  
 QY 1344 CCTCATCTTCAATATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1403  
 Db 2462 CCTCATTTTCAACATCTTTTACTGGATCACAACAAGATCAITTCGGCATGAAGATGTCCA 2521  
 QY 1404 CCAG 1407  
 Db 2522 CAAG 2525

RESULT 7

US-10-486-706-235  
 ; Sequence 235, Application US/10486706  
 ; Publication No. US2005007108A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: LANDFIELD, PHILIP W.  
 ; APPLICANT: BLALOCK, ERIC M.  
 ; APPLICANT: CHEN, KUEY-CHU  
 ; APPLICANT: FOSTER, THOMAS C.  
 ; TITLE OF INVENTION: GENE EXPRESSION PROFILE BIOMARKERS AND THERAPEUTIC TARGETS FOR  
 ; TITLE OF INVENTION: BRAIN AGING AND AGE-RELATED COGNITIVE IMPAIRMENT  
 ; FILE REFERENCE: 50229-426  
 ; CURRENT APPLICATION NUMBER: US/10/486,706  
 ; CURRENT FILING DATE: 2004-02-13  
 ; PRIOR APPLICATION NUMBER: PCT/US02/25607  
 ; PRIOR FILING DATE: 2002-08-13

; PRIOR APPLICATION NUMBER: US 60/311,343  
 ; PRIOR FILING DATE: 2001-08-13  
 ; NUMBER OF SEQ ID NOS: 461  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 235  
 ; LENGTH: 3865  
 ; TYPE: DNA  
 ; ORGANISM: Rattus norvegicus  
 ; US-10-486-706-235  
 Query Match 39.9%; Score 654.8; DB 19; Length 3865;  
 Best Local Similarity 70.8%; Pred. No. 2.9e-199;  
 Matches 938; Conservative 0; Mismatches 332; Indels 54; Gaps 3;  
 QY 84 AAAAGAGGAGTCAAAATCTGGAACCAAGGGGTCCAGCCCATGTCCTCCCTCTGATTTTAAAG 143  
 Db 1256 AGACCATGACTCCAGGTCTGGAACCATCCCTGCAGACCTCTGCTCTTCAGATTTCTT 1315  
 QY 144 AGCAAACTTATGGGGCGAACATCTGGATATGATGCGAGATTCGGGCCCAATTTTAAAG 203  
 Db 1316 GGATAAACTAATGGGAAGGACATCAGGATATGATGCAAGAAATCAGGCCAAATTTTAAAG 1375  
 QY 204 CCACCCGCTGAAGCTGACCTGCAACATCTTCAACACAGTTTTCAGCTCCGTCAACAGAC 263  
 Db 1376 TCCTCAGTAAAGCTTACTTTCGATATTTTATCAACAGTTTGGATCGGTCAAGAAC 1435  
 QY 264 CAAATGGAATACCCGGGTGAATGTCTTTCGGGCAACAGTGGAAATGACCCACGCTGTC 323  
 Db 1436 CACCATGGACTACCGAGTGAACATTTTCTGAGACAGCAGTGGAAACGATTCACGGCTGGC 1495  
 QY 324 CTACCGAAGATATCTCTGATGACTCTCTGGACCTGATCCCTCCATGCTGAGCTCTATCTG 383  
 Db 1496 ATACAGTGAATACCCAGATGATTTCCCTGGATTTGGATTCATCGATGCTGATTTGAT 1555  
 QY 384 GAAGCAGACCTCTCTTCTTAAAGAGAGAGGCGCAACTTCCATGAGTGAACACGGA 443  
 Db 1556 GAACCAAGATTTGTTCTTCCCAATGAGAAAGAGCAATTTCCATGATGTACCACTGA 1615  
 QY 444 CAACAAATGCTGCGCATCTTCAAGAAATGGGAATGTGCTGTACAGCATCAGGCTGACCT 503  
 Db 1616 TAACAGTTGCTGGGATTTTCCAAATGCAAGTGTCTACAGATTTAGACTCACTT 1675  
 QY 504 CATTTGCTGCTGATGAGCTCAAGAACTTCCCATGGAATCCAGACGTGAGGAT 563  
 Db 1676 GACTTTATCTGCTCCATGGACCTGAAGAACTTTTCAATGGATGTCCAGACCTGTCAAT 1735  
 QY 564 GCAGCTTGAAGAGCTCATCTACTCTGCAGCCCTCTGCCATCTCTGCTCACTTTTCAAGTGG 623  
 Db 1736 GCAGCTGGAGAGTT-----TTGG 1753  
 QY 624 CTACACCATGAAGACCTCGTGTGAGTGGCTGGAAGATGCTCTGCTGCTGCTGCTGCTGCTGCT 683  
 Db 1754 GTACACCATGAATGACTGATATTTGAGTGGTAAAGTATGATGCTGCTGCTGCTGCTGCTGCT 1810  
 QY 684 TGAGGGGCTGACTCTGCCCCAGTTTATCTTGGGGATGAGAAGATCTAGGCTGTGTGAC 743  
 Db 1811 TGAAGGACTCACCTGCTCAGTTTATTTTGAAGAAGAGAGAACTTTGGCTATTGTCAC 1870  
 QY 744 CAAGCACTACACACAGGGAAATTCACCTGCAATCGAGGTAAAGTTTCACTCGAAGCGCA 803  
 Db 1871 AAAGCAATTAACACTGGCAAGTTTACCTGCAATGAGGTCAAGTTTCACTCGAAGCGCA 1930  
 QY 804 GATGGGCTACTATCTGATTCAGATGTACATCCCGACCTACTCATCTGCTGCTGCTGCTGCTGCTGCT 863  
 Db 1931 GATGGCTATTATTTGATCCAGATGTATATCCCGACCTGCTGATAGTCAATTTTGTCTGCTG 1990  
 QY 864 GGTCTCCTCTGATCAACATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 923  
 Db 1991 GGTCTCCTTTTGGATAAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2050  
 QY 924 CGTCTCACCATGACCAACCCAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 983  
 Db 2051 CGTCTGCAATGACTACCCAGAGTTCAAGTTTCCAGGGCATCTCTGCCAAAGTCTCATA 2110





QY 980 CCTACGTGAAGCAATCGACATCTGATGGCTGTGTCTGTCTTTGTGCTGCCT 1039  
 Db 1352 CATATGTCAAAGTATATGATTAATTTGGATGGCAGTATGCTCTTTTGTGTTTCAGCAC 1411  
 QY 1040 TGTGTGATGATGCTGCCATAAAATTTTGTCTGTGAGTAAAGAAATTCATACGACTTC 1099  
 Db 1412 TTCTGAGTATGAGCTGTAAATTTTGTATCAAGACACACAAGAACTTCGAGATTTC 1471  
 QY 1100 GAAGAAGGCAAGCGCCCAACGCTTGGAGGAAGATATCATCCAAAGAAAGTCGTTCTATT 1159  
 Db 1472 GACGAAAGAGAAAG-----AATAAGGATGATGAGGTAAAGGAAAGCCGATTTCAGCT 1522  
 QY 1160 TCCGTGGCTATGCTTGGGCCACTGCTGAGGCAAGAGATGAGGTCCATGGAAGGTT 1219  
 Db 1523 TCACAGCCTATGGAATGGGACCATGCTTACAAGCAAGAGGATGCGATGACTCCAAAGGGCC 1582  
 QY 1220 CTGGCATTTATAGTCCCAACCTCCAGCCCTCTTCTTAAGGGAGAGAGAAACACCGGA 1279  
 Db 1583 C-----CAACCCCTCTGTCAGGTATGCAAAAGTCTGATGAATGAGGA 1630  
 QY 1280 AACTCTACGTGACTGAGCAAGAGAAATGACACCATCTCCCGGGCTGTCTTCCCTTTCA 1339  
 Db 1631 AGCTCTTTATCGACCGGGCCCAAGAGATTGATACCATCTCCCGAGCTGCTTCCCATTTAG 1690  
 QY 1340 CTTTCCTCATCTTCAATATCTTCTACTGGTGTCTTATAAGTGTATGTCAGAGATA 1399  
 Db 1691 CTTTTTTGATTTTAAATATTTTCTACTGGGTTATCTATATAAATTTCTTAGGCATGAGGATA 1750  
 QY 1400 TCCACCAG 1407  
 Db 1751 TTTATCAG 1758

RESULT 9

US-10-757-262-55  
 ; Sequence 55, Application US/10757262  
 ; Publication No. US20040197825A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Karicheti, Venkateswarlu  
 ; APPLICANT: Silos-Santiago, Inmaculada  
 ; APPLICANT: Eliaso, Scott D.  
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING  
 ; TITLE OF INVENTION: UROLOGICAL DISORDERS USING 44390, 54181, 211, 5687, 884,  
 ; TITLE OF INVENTION: 1405, 636, 4421, 5410, 30905, 2045, 16405, 18560, 2047,  
 ; TITLE OF INVENTION: 33751, 52872, 14063, 20739, 32544, 43239, 44373, 51164,  
 ; TITLE OF INVENTION: 53010, 16952, 1587, 2207, 22245, 2387, 52908, 69112, 14990,  
 ; TITLE OF INVENTION: 18547, 115, 579, 15985, 15625, 760, 18603, 2395, 2554, 8675,  
 ; TITLE OF INVENTION: 32720, 4809, 14303, 16816, 17827, 32620, 577, 619, 1423,  
 ; TITLE OF INVENTION: 2158, 8263, 15402, 16209, 16386, 21165, 30911, 41897, 1643,  
 ; TITLE OF INVENTION: 2543, 9626, 13231, 32409, 84260, 2882, 8203, 32678 OR  
 ; FILE REFERENCE: MPI03-007PIRNONMIN  
 ; CURRENT APPLICATION NUMBER: US/10757, 262  
 ; CURRENT FILING DATE: 2004-01-14  
 ; PRIOR APPLICATION NUMBER: US 60/440, 318  
 ; PRIOR FILING DATE: 2003-01-15  
 ; PRIOR APPLICATION NUMBER: US 60/444, 783  
 ; PRIOR FILING DATE: 2003-02-04  
 ; PRIOR APPLICATION NUMBER: US 60/457, 901  
 ; PRIOR FILING DATE: 2003-03-27  
 ; PRIOR APPLICATION NUMBER: US 60/468, 775  
 ; PRIOR FILING DATE: 2003-05-08  
 ; PRIOR APPLICATION NUMBER: US 60/471, 614  
 ; PRIOR FILING DATE: 2003-05-19  
 ; PRIOR APPLICATION NUMBER: US 60/478, 742  
 ; PRIOR FILING DATE: 2003-06-16  
 ; PRIOR APPLICATION NUMBER: US 60/488, 529  
 ; PRIOR FILING DATE: 2003-07-18  
 ; PRIOR APPLICATION NUMBER: US 60/491, 156  
 ; PRIOR FILING DATE: 2003-07-30  
 ; PRIOR APPLICATION NUMBER: US 60/499, 594  
 ; PRIOR FILING DATE: 2003-09-02

; PRIOR APPLICATION NUMBER: US 60/506, 332  
 ; PRIOR FILING DATE: 2003-09-26  
 ; NUMBER OF SEQ ID NOS: 136  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 55  
 ; LENGTH: 3069  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (421)...(1770)  
 ; US-10-757-262-55

Query Match 33.7%; Score 552.8; DB 18; Length 3069;  
 Best Local Similarity 66.6%; Pred. No. 1.9e-166;  
 Matches 885; Conservative 0; Mismatches 377; Indels 66; Gaps 4;  
 QY 80 TGGCAAAAGAGGAAGTCAAAATCTGGAAACCAAGGGGTCACAGCCCATGTCCCTCTGTGATT 139  
 Db 497 TTGCCCAAGGAACAGACAGTGCAGATCTCGAAGTGTCTCAATGTCACTTCTGATT 556  
 QY 140 TCCTAGACAAACTTATGGGGCGAACAATCTGGATATGATGCGAGATTCGGCCCAATTTTA 199  
 Db 557 TTCTGATTAATTAATGGGCGAGACATCAGGATATGATCAAGAAATCAGACCCCAATTTTA 616  
 QY 200 AAGGCCACCCGTGAACGTGACCTGCAACATCTTCAACAGTTTTCAGTCCGTACCA 259  
 Db 617 AAGGCCCTCCAGTTAATGTGCATGCAACATATTTCAACAGTTTTCGGCTCTATCGCAG 676  
 QY 260 AGACCAACATGACATACCGGTGAATGTCTTTGGGGCAACAGTGGAAATGACCCAGGCC 319  
 Db 677 AGACGACCATGATTAACAGATGATATCTTTCTCGTGAAGATGATGATCCCGGCC 736  
 QY 320 TGTCTTACCGAGAAATATCTGTGATGACTCTCTGGACCTCGATCCCTCCATCTGACTCTA 379  
 Db 737 TCGGTACAGTGAATCTGTGAGACTCTTTAGACCTCGACCTCCATGTTGACTCCA 796  
 QY 380 TGTGAAGCGACACTCTTTCTTGTAAATGAGAAAGGGGCAACTTCCATGAGGTGACCA 439  
 Db 797 TTTGGAAACCTGATTTGTTCTTTGCAATGAAAGGGTCCCAACTTTCATGAAGTCACTA 856  
 QY 440 CGGACAAACAGTTACTCGGCATCTTCAAGATGGGAATGTCTGTACAGCATCAGGCTGA 499  
 Db 857 CAGACAAACAAATGCTGAAGAAATTTTCAAAATGGAATGTTCTTTATCAATGAATTA 916  
 QY 500 CCTCATTTTGTCTCTGCTGATGAGACTCAAGAACTTCCCATGAGCATCCAGACGTGCA 559  
 Db 917 CATTAACACTTTCTGCTCAATGGATCTCAAGAAATTTTCCCATGGATGTACAACATGTA 976  
 QY 560 CGATGAGCTTGAGAGCTCATCCATCTGTGACGCCCTCTGCCCATCTCTGTCTACTTTTCA 619  
 Db 977 TAATGCAACTGGAAAGCT----- 994  
 QY 620 TTGGCTACACCATGAAGACCTCGTCTTTGAGTGGCTGGAGATGCTCTGTGTCCAAG 679  
 Db 995 TTGGGTACAAATGAATGATCTCATTTTGAATGGC---AAGATGAGGCCACCGTCAAG 1051  
 QY 680 TGGCTGAGGGGCTGACTCTGCCCCAGTTTATCTTGGGGATGAGAAGGATCTAGGCTGTT 739  
 Db 1052 TGGCAGAGGACTCACTTTTGCCCACTTTCTGTTGAAGAGAAAGATTTACGTACT 1111  
 QY 740 GTACCAAGCACTACACACAGGAAATTCACCTGCAATCGAGGTAAAGTTTCACTGGAAC 799  
 Db 1112 GCACTAAACATTAACAATACAGGAAAGTTTACGTGTATAGAAGTGCAGATTCATCTGAGC 1171  
 QY 800 GGCAGATGGCTACTATCTGATTCAGATTCAGATCCCGACCTACTCATCTGCTCTCTGT 859  
 Db 1172 GACAAATGGGACTACTATCTGATCCAGATGTACATCCCACTCTCTGATTTTATCTAT 1231  
 QY 860 CTTGGGTCTCTTCTGATCAACATGGATGTGCTGCCCTCTGCCCTGCTGGGCTGCGGATCA 919  
 Db 1232 CTTGGGTTTCTGTTCTGATCAACATGATGACGACCGGCCAGGTAGCTCTGGGNTAA 1291

QY 920 CCACCGTGTACCATGACACCCAGAGCTCTGGCTCCGGGCTCTTTGCTTAAGGTG 979  
 Db 1292 CCACCGTGTACCATGACATACAGAGTTTCAGGATCACGAGCTTCTTGCCCAAAAGTTT 1351  
 QY 980 CCTAGTGAAGGAATCGACATCTGATGGCTGTGTCTGTCTTTGTTGGCTGCCT 1039  
 Db 1352 CATATGTCAAAGCTATTGATTAATTGGATGGCAGTATGCTCTTTTGTGTTTTCAGCAC 1411  
 QY 1040 TGGTGGAGTATGCTGCGCATAAATTTTGTCTCTGTCAGCATAAAGAAATCATACGACTTC 1099  
 Db 1412 TTCTGGAGTATGAGCTGTAAATTTTGTATCAAGACAACAAGAACTTCTGAGATTTC 1471  
 QY 1100 GAAGAAGGCAGAGCGCCACGCTTGGAGGAAGATATCATCCAGAAAGTCTGTTCTATT 1159  
 Db 1472 GACGAAGAGAAAG-----AATAAGGATGATGAGGTAAAGGAAGCCGATTCAGCT 1522  
 QY 1160 TCCGTGGCTATGGCTTGGGCCATCTGCTGAGGCAAGAGATGAGGTCCAATGGAAGGTT 1219  
 Db 1523 TCACAGCCTATGGAATGGGACCATGTCTACAAGCAAAAGGATGGCATCACTCCAAAGGGCC 1582  
 QY 1220 CTGGCATTTATAGTCCCCCACTCCAGCCCTCTTCTAAGGGAAGGAGAAACCCAGCGGA 1279  
 Db 1583 C-----CAACCCCTCTGTCAGGTAATGCCAAAGTCTCTGATGAAATGAGGA 1630  
 QY 1280 AACTCTAGTGGAGTCTGAGGCAAGAGAAATTGACACCATCTCCGGGCTGCTTCCCTTTCA 1339  
 Db 1631 AGGTCTTTATCGACCGGGCCAGAGATTGATACCATCTCCCGAGCTGCTTCCCATAG 1690  
 QY 1340 CTTTCTCATCTTCAATATCTTCTACTGGTGTCTTATAAGTGTATGTCAGAGATA 1399  
 Db 1691 CTTTTTGTATTTTAAATATTTCTACTGGGTATCTATAAAATTTCTAGGCATGAGGATA 1750  
 QY 1400 TCCACAG 1407  
 Db 1751 TCCATCAG 1758

RESULT 10  
 US-10-075-846-5  
 ; Sequence 5, Application US/10075846  
 ; Publication No. US20030032608A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bristol-Myers Squibb Company  
 ; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING A NOVEL GLYCINE RECEPTOR ALPHA SUBUNIT B  
 ; TITLE OF INVENTION: IN THE GASTROINTESTINAL TRACT, HGRA4, and SPLICE VARIANT THEREOF  
 ; FILE REFERENCE: D0079 NP  
 ; CURRENT APPLICATION NUMBER: US/10/075,846  
 ; CURRENT FILING DATE: 2002-02-13  
 ; PRIOR APPLICATION NUMBER: US 60/269,535  
 ; PRIOR FILING DATE: 2001-02-16  
 ; NUMBER OF SEQ ID NOS: 81  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 5  
 ; LENGTH: 679  
 ; TYPE: DNA  
 ; ORGANISM: homo sapiens  
 US-10-075-846-5

Query Match 29.3%; Score 480.6; DB 14; Length 679;  
 Best Local Similarity 99.2%; Pred. No. 1.6e-143;  
 Matches 483; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1118 AACGCTTGGAGGAAGATATCATCCAGAAAGTCTTTCTATTTCCTGGCTATGGCTTG 1177  
 Db 193 AATTCAGGAGGAAGATATCATCCAGAAAGTCTTTCTATTTCCTGGCTATGGCTTG 252  
 QY 1178 GCACCTGCTTCGAGGCAAGAGATGGAGTCCAATGGAAGGTTCTGGCAITTTATAGTCCCC 1237  
 Db 253 GCACCTGCTTCGAGGCAAGAGATGGAGTCCAATGGAAGGTTCTGGCAITTTATAGTCCCC 312  
 QY 1238 AACCTCAGGCCCTCTTCTAAGGGAAGGAGAAACCCAGCGGAACTCTACGTGAGCTGAG 1297  
 Db 313 AACCTCAGGCCCTCTTCTAAGGGAAGGAGAAACCCAGCGGAACTCTACGTGAGCTGAG 372

QY 1298 CCAAGAGAATTGACACCATCTCCCGGGCTGTCTTCCCTTTCACTTTCTCTCAATA 1357  
 Db 373 CCAGAGAAATTGACACCATCTCCCGGGCTGTCTTCCCTTTCACTTTCTCTCAATA 432  
 QY 1358 TCTTCTACTGGGTGTCTATAAAGTGTATGTCAGAGATATCCACAGGCTCTGTGAA 1417  
 Db 433 TCTTCTACTGGGTGTCTATAAAGTGTATGTCAGAGATATCCACAGGCTCTGTGAA 492  
 QY 1418 TAGGTGGGAGCTATAGAGTCTGCTGCTGGGCTCTCTGCTTCTCTGGTGGGCTTCT 1477  
 Db 493 TAGGTGGGAGCTATAGAGTCTGCTGCTGGGCTCTCTGCTTCTCTGGTGGGCTTCT 552  
 QY 1478 CCTCAGTTTAGACTCCATTAGGGGTTTGACACAGTTCCTTCTCTGATCTCCCACTCAGAAT 1537  
 Db 553 CCTCAGTTTAGACTCCATTAGGGGTTTGACACAGTTCCTTCTCTGATCTCCCACTCAGAAT 612  
 QY 1538 TCAACTACCACTCCCAAGCTATGTTGGGCTATATTCATGATGTCGCAATGGTGGCTGTAC 1597  
 Db 613 TCAACTACCACTCCCAAGCTATGTTGGGCTATATTCATGATGTCGCAATGGTGGCTGTAC 672  
 QY 1598 TTATAAA 1604  
 Db 673 TTATAAA 679

RESULT 11  
 US-09-802-668-49/c  
 ; Sequence 49, Application US/09802668  
 ; Publication No. US20030190714A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wood, Linda  
 ; APPLICANT: Vogel, Gabriel  
 ; APPLICANT: Karnovsky, Alla  
 ; APPLICANT: Linske-O'Connell, Lisa I.  
 ; APPLICANT: Wang, Jun  
 ; APPLICANT: Liu, Derong  
 ; TITLE OF INVENTION: Human Ion Channels  
 ; FILE REFERENCE: 00069US1  
 ; CURRENT APPLICATION NUMBER: US/09/802,668  
 ; CURRENT FILING DATE: 2001-03-09  
 ; PRIOR APPLICATION NUMBER: 60/188,517  
 ; PRIOR FILING DATE: 2000-03-10  
 ; NUMBER OF SEQ ID NOS: 117  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 49  
 ; LENGTH: 554  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-802-668-49

Query Match 18.2%; Score 299; DB 10; Length 554;  
 Best Local Similarity 98.1%; Pred. No. 4.9e-85;  
 Matches 313; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1118 AACGCTTGGAGGAAGATATCATCCAGAAAGTCTTTCTATTTCCTGGCTATGGCTTG 1177  
 Db 319 AATTCAGGAGGAAGATATCATCCAGAAAGTCTTTCTATTTCCTGGCTATGGCTTG 260  
 QY 1178 GCACCTGCTTCGAGGCAAGAGATGGAGTCCAATGGAAGGTTCTGGCAITTTATAGTCCCC 1237  
 Db 259 GCACCTGCTTCGAGGCAAGAGATGGAGTCCAATGGAAGGTTCTGGCAITTTATAGTCCCC 200  
 QY 1238 AACCTCCAGCCCTCTTCTAAGGGAAGGAGAAACCCAGCGGAACTCTACGTGAGCTGAG 1297  
 Db 199 AACCTCCAGCCCTCTTCTAAGGGAAGGAGAAACCCAGCGGAACTCTACGTGAGCTGAG 140  
 QY 1298 CCAAGAGAATTGACACCATCTCCCGGGCTGTCTTCCCTTTCACTTTCTCTCAATA 1357  
 Db 139 CCAGAGAAATTGACACCATCTCCCGGGCTGTCTTCCCTTTCACTTTCTCTCAATA 80  
 QY 1358 TCTTCTACTGGGTGTCTATAAAGTGTATGTCAGAGATATCCACAGGCTCTGTGTA 1416

```

Db 79 TCTTCTACTGGTGTCTATAAAGTGTCTACGGTCTAGAGATATCCACAGGCTTCTGTGA 20
QY 1417 ATAGGGTGGAGCTATAGA 1435
Db 19 ATAGGGTGGAGCTATAGA 1
RESULT 12
US-10-243-475-49/c
; Sequence 49, Application US/10243475
; Publication No. US20030194720A1
; GENERAL INFORMATION:
; APPLICANT: Roberds, Steven L.
; APPLICANT: Benjamin, Christopher W.
; APPLICANT: Karnovsky, Alla M.
; APPLICANT: Ruble, Cara L.
; TITLE OF INVENTION: Human Ion Channels
; FILE REFERENCE: PHR0039-100
; CURRENT APPLICATION NUMBER: US/10/243,475
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: 60/403,254
; PRIOR FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: 60/318,733
; PRIOR FILING DATE: 2001-09-12
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 49
; LENGTH: 554
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-243-475-49
Query Match 18.2%; Score 299; DB 16; Length 554;
Best Local Similarity 98.1%; Pred. No. 4.9e-85;
Matches 313; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
QY 1118 AACGCTGGAGGATATCATCAAGAAAGTGTCTTATTTCCGTGGCTATGGCTTGG 1177
Db 319 AATTCAGGAGGAGATATCATCAAGAAAGTGTCTTATTTCCGTGGCTATGGCTTGG 260
QY 1178 GCCACTGCCCTCGAGGAGAGATGGAGTCCATGGAAGTGTCTGGCAATTTATAGTCCCC 1237
Db 259 GCCACTGCCCTCGAGGAGAGATGGAGTCCATGGAAGTGTCTGGCAATTTATAGTCCCC 200
QY 1238 AACCTCCAGCCCTCTTCTTAAGGAGAGAGAAACACGCGGAACTCTACGTGGAGTGG 1297
Db 199 AACCTCCAGCCCTCTTCTTAAGGAGAGAGAAACACGCGGAACTCTACGTGGAGTGG 140
QY 1298 CCAAGAGATTTGACCATCTCCCGGGCTGTCTTCCCTTTCACTTTCCCTCATCTTCAATA 1357
Db 139 CCAAGAGATTTGACCATCTCCCGGGCTGTCTTCCCTTTCACTTTCCCTCATCTTCAATA 80
QY 1358 TCTTCTACTGGTGTCTATAAAGTGTCTATGGTCTAGAGATATCCACAGGCTTCTGTGA 1416
Db 79 TCTTCTACTGGTGTCTATAAAGTGTCTATGGTCTAGAGATATCCACAGGCTTCTGTGA 20
QY 1417 ATAGGGTGGAGCTATAGA 1435
Db 19 ATAGGGTGGAGCTATAGA 1
RESULT 13
US-10-305-720-271
; Sequence 271, Application US/10305720
; Publication No. US20040010136A1
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice K.; Seilhamer, Jeffrey J.
; TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expression
; FILE REFERENCE: PA-0002-1 CON
; CURRENT APPLICATION NUMBER: US/10/305,720
; CURRENT FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: 09/016,434
; PRIOR FILING DATE: 1998-01-30

```

```

; NUMBER OF SEQ ID NOS: 1490
; SOFTWARE: PERL Program
; SEQ ID NO 271
; LENGTH: 253
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040010136A1 1695770
US-10-305-720-271
Query Match 15.4%; Score 253; DB 17; Length 253;
Best Local Similarity 100.0%; Pred. No. 2.1e-70;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1004 GGATGGCTGTGTCTGTCTCTTTTGTGTTTCGCTGCTTGTGGAGTATGTCGCATAAATT 1063
Db 1 GGATGGCTGTGTCTGTCTCTTTTGTGTTTCGCTGCTTGTGGAGTATGTCGCATAAATT 60
QY 1064 TTGTTTCTCGTCAGCATAAAGAAATTCATACGACTTCGAAGAGGCGACGCGCAACGCT 1123
Db 61 TTGTTTCTCGTCAGCATAAAGAAATTCATACGACTTCGAAGAGGCGACGCGCAACGCT 120
QY 1124 TGGAGGAAGATATCATCAAGAAAGTCGTTTCTATTTCCGTGGCTATGGCTTGGGCACT 1183
Db 121 TGGAGGAAGATATCATCAAGAAAGTCGTTTCTATTTCCGTGGCTATGGCTTGGGCACT 180
QY 1184 GCCTGCAGCAAGAGATGGAGTCCCAATGGAAGTTCTGGCAATTTATAGTCCCAACCTC 1243
Db 181 GCCTGCAGCAAGAGATGGAGTCCCAATGGAAGTTCTGGCAATTTATAGTCCCAACCTC 240
QY 1244 CAGCCCTCTTCT 1256
Db 241 CAGCCCTCTTCT 253
RESULT 14
US-10-172-118-529
; Sequence 529, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 529
; LENGTH: 2106
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_000824
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-529
Query Match 14.2%; Score 232.6; DB 17; Length 2106;
Best Local Similarity 56.8%; Pred. No. 2.6e-63;
Matches 504; Conservative 0; Mismatches 339; Indels 48; Gaps 2;
QY 170 GATATGATCCAGGATTCGGCCCAATTTTAAAGGCCACCGGTGAACGTGACCTGCAACA 229
Db 273 GTTATGATCCAGGATTAAGCAACTTCAGAGGCATTCCTGTTGATAGTAGTCAACA 332

```

QY 230 TCTTCATCAACAGTTTCAGCTCCGTCACCAAGACCAATGGACTACCGGTGAATGTCT 289  
Db 333 TTTTATTAACAGTTTGGATCCATTCAAGAAACAATGGACTATAGATTAAATCT 392  
QY 290 TCTTGGGCAACAGTGAATGACCCAGCGCTGCTCCTACCGAGATAT-----CCTGATG 343  
Db 393 TCTGAGACAAAAATGGAATGACCCAGCGCTGAAGCTCCCAAGTATTTTGGGGTTCAG 452  
QY 344 ACTCTCTGGACCTCGATCCCTCATCTGAGCTCTATCTGGAAGCAGACCTCTCTCTTG 403  
Db 453 ATGCACTGACAGTGGATCCCAATTTTATGATGTGACCACTTCAAGGCAATCTCTCTTATTT 512  
QY 404 CTAATGAGAAAGGGGCAACTTCCATGAGTGACCAAGCAACAAGTTACTGCGCATCT 463  
Db 513 CAAATGAAAAAGTGCATTTTTCATGATGTGACCCAGGAAACATCTCTCTTATTT 572  
QY 464 TCAAGATGGGAATGTCTGTACAGATCAGGCTGACCCCTCATTTTGTCTGCTGATGG 523  
Db 573 TTCTGTGATGAGATGCTCTTGTGAGATGAGTTATCTATTCTCTTTCATGCTTGG 632  
QY 524 ACCTCAAGAACTTCCCATGAGATCCAGACGTCGACGATGAGCTTGAAGCTCATCCA 583  
Db 633 ACTTGACATGTTTCCATGGATACAAAGTTGCAAGTGAAGTGAAGCT-----686  
QY 584 TACTCTGACGCTCTGCCATCTCTCTCACTTTTCACTTGGCTACACCATGAAGACCTCG 643  
Db 687 -----TTGTTTACCAACTGATTTAC 710  
QY 644 TGTTTGAAGTGGTGAAGATGCTCTGCTGTCCAAAGTGGCTGAGGGCTGACTCTGCCCC 703  
Db 711 GATTTATCTGGCAGTCAAGAGATCTGTGCAATTTAGAAAAATTTGCTTGCCTCAATTTG 770  
QY 704 AGTTTATCTTGGGGATGAGAGGATCTAGGCTGTGTACCAAGCACTACAAACACAGGGA 763  
Db 771 ATATCAAAAGGAAGATATTGAATATGTTAACTGTACAAAAATATAAAGGCAACGGCT 830  
QY 764 AATTCACCTGATCGAGGTAAAGTTTCACTGGAACGGCAGATGGGCTACTATCTGATTC 823  
Db 831 ACTACACATGCTGGAAAGTCATCTTCCACCTGAGGAGGAGGTGCGCTTTTACATGATGG 890  
QY 824 AGATGTACATCCCGCTTACTCATCTCTGCTCTGCTCTCTCTCTCTCTCTCTCTCTCT 883  
Db 891 GGGTCTACCCCAACCCCTGCTCATTTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 950  
QY 884 TGGATGCTGCCCTGCTGCCCTGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 943  
Db 951 CGGACGCGAGTGTGCGCAGAGTCCCTGGGTATCTTCTCAGTCTCTCAGTCTCTGCTCTG 1010  
QY 944 AGAGCTCTGGCTCCCGGGCTCTTTGGCTTAAGTGTCTTGTGCTGCTGCTGCTGCTGCT 1003  
Db 1011 AGTGCAACCCCTGCGCTGAGCTTCCCAAGTTTCCATATGTGAAGGCTCTTGATGTTT 1070  
QY 1004 GATGCTGTGTGTCTGCTTGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1054  
Db 1071 GGCTTATGCTGCTCTCTCTTGGGTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1121

RESULT 15

US-10-342-887-529  
; Sequence 529, Application US/10342887  
; Publication No. US20040058340A1  
; GENERAL INFORMATION:  
; APPLICANT: Dai, Hongyue  
; APPLICANT: He, Yudong  
; APPLICANT: Linsley, Peter S.  
; APPLICANT: Mao, Mao  
; APPLICANT: Roberts, Christopher J.  
; APPLICANT: Van 't Veer, Laura Johanna  
; APPLICANT: Van de Vijver, Marc J.  
; APPLICANT: Bernards, Rene  
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients  
; FILE REFERENCE: 9301-188-999  
; CURRENT APPLICATION NUMBER: US/10/342,887

; CURRENT FILING DATE: 2003-01-15  
; PRIOR APPLICATION NUMBER: 60/298,918  
; PRIOR FILING DATE: 2001-06-18  
; PRIOR APPLICATION NUMBER: 60/380,710  
; PRIOR FILING DATE: 2002-05-14  
; PRIOR APPLICATION NUMBER: 10/172,118  
; PRIOR FILING DATE: 2002-06-14  
; NUMBER OF SEQ ID NOS: 2699  
; SEQ ID NO 529  
; LENGTH: 2106  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; US-10-342-887-529  
  
Query Match 14.2%; Score 232.6; DB 17; Length 2106;  
Best Local Similarity 56.6%; Pred. No. 2.6e-63;  
.Matches 504; Conservative 0; Mismatches 339; Indels 48; Gaps 2;  
  
QY 170 GATATGATCCAGGATTCGGCCCAATTTTAAAGGCCACCCGTCGAACTGACCTGCAACA 229  
Db 273 GTTATGATCCAGGATAAGACCAAACTTCAAGGCAATCTCTGTGATGTAGTAGTCAACA 332  
QY 230 TCTTCAATCAACAGTTTCAGCTCCGTCAACAAGCAACAATGGAATGACCTACCGGTGAATGTCT 289  
Db 333 TTTTATTAACAGTTTGGATCCATTCAAGAAACAACAATGGAATGATAGATTAACATCT 392  
QY 290 TCTTGGGCAACAGTGGATGACCCAGCGCTCTCTTACCGAGAAATAT-----CCTGATG 343  
Db 393 TCTGAGACAAAAATGGAATGACCCAGGCTGAAGCTCCCAAGTGAATTTTAGGGTTCAG 452  
QY 344 ACTCTCTGACCTCGATCCCTCATCTGCTGACTCTATCTGGAAGCCAGACCTCTCTCTTTG 403  
Db 453 ATGCACTGACAGTGGATCCCAATTTTATGATGTGACCACTTATGGAACCTGATTTATTTT 512  
QY 404 CTAATGAGAAAGGGGCAACTTCCATGAGTGACCAAGCAACAAGTTACTGCGCATCT 463  
Db 513 CAAATGAAAAAGTGCCTATTTTTCATGATGTGACCCAGGAAACATCTCTCTCTTATTT 572  
QY 464 TCAAGATGGGAATGTCTGTACAGCATCAGCTGACCTCATTTTGTCTCTGCTGATGG 523  
Db 573 TCTGATGAGATGCTCTTGTGATGAGTTATCTATTACTCTTTCATGCTTGG 632  
QY 524 ACCTCAAGAACTTCCCATGAGATCCAGACGTCGACGATGCAAGTGTGAGAGCTCATCCA 583  
Db 633 ACTTGACATGTTTCCATGGATACAAAGTTGCAAGTGAAGTGAAGCT-----686  
QY 584 TACTCTGACGCTCTGCCATCTCTGTCACTTTTCACTTGGCTACACCATGAAGACCTCG 643  
Db 687 -----TTGTTTACCAACTGATTTAC 710  
QY 644 TGTTTGAAGTGGTGAAGATGCTCTGCTGTCCAAAGTGGCTGAGGGCTGACTCTGCCCC 703  
Db 711 GATTTATCTGGCAGTCAAGAGATCTGTGCAATTTAGAAAAATTTGCTTGCCTCAATTTG 770  
QY 704 AGTTTATCTTGGGGATGAGAGGATCTAGGCTGTGTACCAAGCACTACAAACACAGGGA 763  
Db 771 ATATCAAAAGGAAGATATTGAATATGTTAACTGTACAAAAATATAAAGGCAACGGCT 830  
QY 764 AATTCACCTGATCGAGGTAAAGTTTCACTGGAACGGCAGATGGGCTACTATCTGATTC 823  
Db 831 ACTACACATGCTGGAAAGTCATCTTCCACCTGAGGAGGAGGTGCGCTTTTACATGATGG 890  
QY 824 AGATGTACATCCCGCTTACTCATCTCTGCTCTGCTCTCTCTCTCTCTCTCTCTCTCT 883  
Db 891 GGGTCTACCCCAACCCCTGCTCATTTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 950  
QY 884 TGGATGCTGCCCTGCTGCCCTGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 943  
Db 951 CGGACGCGAGTGTGCGCAGAGTCCCTGGGTATCTTCTCAGTCTCTCAGTCTCTGCTCTG 1010  
QY 944 AGAGCTCTGGCTCCCGGGCTCTTTGGCTTAAGTGTCTTGTGCTGCTGCTGCTGCTGCT 1003  
Db 1011 AGTGCAACCCCTGCGCTGAGCTTCCCAAGTTTCCATATGTGAAGGCTCTTGATGTTT 1070



**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 28, 2005, 03:05:59 ; Search time 6689.22 Seconds  
(without alignments)

9332.242 Million cell updates/sec

Title: US-10-075-846-3

Perfect score: 1640

Sequence: 1 agcaactctgttcctgc.....taaaaaaaaaaaaaaaaaa 1640

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: gb\_est1:\*

2: gb\_est2:\*

3: gb\_hic:\*

4: gb\_est3:\*

5: gb\_est4:\*

6: gb\_est5:\*

7: gb\_est6:\*

8: gb\_gss1:\*

9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	740	45.1	1551	3	AK085727 Mus muscu
2	656	40.0	1294	9	AY418494 Homo sapi
3	654.6	39.9	1294	9	AY418496 Mus muscu
4	640.6	39.1	2387	3	AK078258 Mus muscu
5	638.2	38.9	3220	3	BC045547 Homo sapi
6	630.8	38.5	1294	9	AY418495 Homo sapi
7	561.8	34.3	1080	9	AY398900 Mus muscu
8	557.2	34.0	1080	9	AY398898 Homo sapi
9	477	29.1	1398	9	AY419225 Mus muscu
10	467	28.5	1398	9	AY419223 Homo sapi
11	458.8	28.0	1382	9	AY419224 Pan trogl
12	442.4	27.0	773	6	CD354404
13	401.8	24.5	662	1	AU169868
14	397	24.2	805	6	CD353729
15	378.6	23.1	540	7	CN266978
16	364.2	22.2	786	7	CO395810
17	346	21.1	760	7	CO421128
18	336.8	20.5	622	6	CD354268
19	311.4	19.0	821	7	CK354120
20	302.4	18.4	817	4	BG404477
21	295.6	18.0	678	1	AV729257
22	292.4	17.8	779	5	BQ43713
23	287.2	17.5	715	9	AY398899
24	274	16.7	624	7	CV024684

25	270	16.5	397	7	CN411646
26	266.6	16.3	922	2	BE981841
27	255.2	15.6	720	2	BE981459
28	239.6	14.6	552	4	BM440664
29	219.4	13.4	2875	3	AK083251
30	219.4	13.4	2991	3	BC027094
31	219.4	13.4	3008	3	BC037605
32	219	13.4	2593	3	BC022502
33	215.6	13.1	536	2	BF906462
34	211	12.9	3586	3	AK038618
35	209.4	12.8	2946	3	AK082578
36	203.8	12.4	895	6	CD301151
37	197.6	12.0	762	4	BI736322
38	193.8	11.8	1398	9	AY416040
39	190.6	11.6	1398	9	AY416042
40	185	11.3	1398	9	AY416041
41	178.2	10.9	1422	9	AY416045
42	177.4	10.8	1398	6	CD013911
43	176.2	10.7	1422	9	AY416043
44	175.6	10.7	1422	9	AY416044
45	175.4	10.7	937	7	CF408510

#### ALIGNMENTS

RESULT 1

AK085727

LOCUS

DEFINITION

AK085727

1551 bp mRNA linear

Mus musculus 10 days lactation, adult female mammary gland cDNA,

RIKEN full-length enriched library, clone:D730029F11

Product:glycine receptor, alpha 4 subunit, full insert sequence.

ACCESSION

AK085727.1

VERSION

HTC; CAP trapper.

KEYWORDS

Mus musculus (house mouse)

SOURCE

Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1

Carninci, P. and Hayashizaki, Y.

AUTHORS

High-efficiency full-length cDNA cloning

JOURNAL

Meth. Enzymol. 303, 19-44 (1999)

MEDLINE

99279253

PUBMED

10349636

REFERENCE

2

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

PUBMED

11042159

REFERENCE

3

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,

Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,

Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,

Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,

Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,

Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,

Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer

Genome Res. 10 (11), 1757-1771 (2000)

PUBMED

11076861

REFERENCE

4

The RIKEN Genome Exploration Research Group Phase II Team and the

FANTOM Consortium.

Functional annotation of a full-length mouse cDNA collection

Nature 409, 685-690 (2001)

REFERENCE

5

The FANTOM Consortium and the RIKEN Genome Exploration Research



```

REFERENCE
AUTHORS
Mammalia: Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 1294)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
JOURNAL
PUBMED
2 (bases 1 to 1294)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
These sequences were made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
Location/Qualifiers
1..1294
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>1294
/genes="GLRAL"
/locus_tag="HCM6576"
ORIGIN
Query Match 40.0%; Score 656; DB 9; Length 1294;
Best Local Similarity 71.9%; Pred. No. 8.3e-172;
Matches 952; Conservative 0; Mismatches 315; Indels 57; Gaps 5;
QY 89 AGGAAGTCAAAATCTGGAACCAAGGGTCCAGCCCATGTCCCTCTGATTTCTCTAGACA 148
DB 15 AGAGAGCTGAGTGTGCTGCGCACCACCAAGCTATGTACCTCGATTTCTGATA 74
QY 149 AACTTATGGGGGAAACATCTGGATATGATGCCAGGATTCGGCCCAATTTTAAAGGCCAC 208
DB 75 AGCTAATGGGGAGAACCTCCGGATATGATGCCAGGATCAGGCCCAATTTTAAAGTCCCC 134
QY 209 CGGTGAACGTGACCTGCACATCTTATCAACAGTTTCAGTCCGTCCGTACCCAGACCA 268
DB 135 CAGTGAACGTGACCTGCACATTTTATCAACAGTCTTGGTTCCATTCGTGAGACCA 194
QY 269 TGGACTACGGGTGAATGTTCTTTCGGCAACAGTGGAAATGACCCAGCTCTCTTACC 328
DB 195 TGGACTATAGGGTCAACATCTTCTCGGGCAGCAATGGAAACGACCCCGCTGGCTATA 254
QY 329 GAGAATATCTGTGATGACTCTCTGGACCTCGATCCCTCCATGCTGGACTCTATCTGGAAGC 388
DB 255 ATGAATACCTTGACGACTCTCTGGACCTGGACCCATCCATCTGGAATCTCGAAAC 314
QY 389 CAGACCTCTCTTGTGTAATGAGAAAGGGGCCAACTTCCATGAGGTGACACACGGAACA 448
DB 315 CTGACCTGTTCTTTGGCAACAGAGAAGGGGCCCACTTCCATGAGATCACCACAGACA 374
QY 449 AGTTACTGGCCATCTTCAAGAATGGGAATGCTGTGTACAGCATCAGCTGACCTCATTT 508
DB 375 AATTGCTAAGGAATCTCCGGAATGGGAATGCTCTTACAGCATCAGATCACCCTGACAC 434
QY 509 TGTCTGCTGATGGACCTCAAGAACTTCCCAATGGACATCCAGAGCTGCACGATGACG 568
DB 435 TGGCTGCCCCATGGACTTGAAGATTTCCCAATGGATGTCAGACATGATATCATGCAAC 494
QY 569 TTGAGAGCTCATCCATCTCTGACGCCCTCTGCCATCTCTGTCACTTTTCAGTTGGCTACA 628
DB 495 TGGAAAGCT-----TTGGATATA 512
QY 629 CCATGAAGACCTCGTGTGTTGAGTGGTGAAGATGCTCTGCTGCTCCAGTGGCTGAGG 688
DB 513 CGATGAATACCTCATCTTTGAGTGGCAGGAACAGG---GAGCCGTGACGTAGCAGATG 569

```

```

QY 689 GSCTGACTCTGCCCCAGTTTATCTTTCGGGATGAGAGGATCTAGGCTGTGTGTACCAAGC 748
DB 570 GACTAACTCTGCCCCAGTTTATCTTGAAGGAAGAGAGGACTTGAATATCTTGACCACAGC 629
QY 749 ACTACAACACACGGGAAATTCACCTGCATCGAGTAAAGTTTACCTTGGAAACGGCAGATGG 808
DB 630 ACTACACACACAGTAAATTCACCTGCATGAGGCCCGGTTCCACCTGGAGCGGCAGATGG 689
QY 809 GTTACTATCTGATTCAGATGATCATCCCGAGCTACTCATGTCGTCTCTGTTCTGTTCT 868
DB 690 GTTACTATCTGATTCAGATGATTAATCCAGGCTGCTCATTTGTCTCTCTCATGATCT 749
QY 869 CTTCTGGATCAACATGATGCTGCCCTGCCCTGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 928
DB 750 CTTCTGGATCAACATGATGCTGCCCTGCCCTGCTGGGCTGGGCTGGGCTGGGCTGGGCT 809
QY 929 TCACCATGACACACCCAGAGCTCTGGCTCCCGGCGCTCTTTTGGCTAAGGTGTCTTACGTGA 988
DB 810 TCACCATGACACACCCAGAGCTCGGCTCTCGAGCATCTCTGCCAAGGTGTCTTATGTA 869
QY 989 AGCAATCGACATCTGATGCTGTGTCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1048
DB 870 AAGCCATTGACATTTGGATGGCAGTTTGCCTGCTCTTGTGTGTGTGTGTGTGTGTGTGT 929
QY 1049 ATGCTGCCATAAATTTTGTCTCTGTCAGCATAAAGAAATTCATACACTTCGAGAGGCG 1108
DB 930 ATGCTGCCATAAATTTTGTCTCTGCGCAACATAAGGAGCTGCTCCGATTCAGAGGAAGC 989
QY 1109 AGAGGCGCCAAACGCTTGGAGGAAGATATCATCCAGAAAGTCTGTTTCTATTTCCTGGCT 1168
DB 990 GGAGACATCAACAGGAGGATGAAGCT-----GGAGAGGCGCTTTTAACTTCTGCTCT 1043
QY 1169 ATGGCTTGGGCC---ACTGCTCGAGGCAAGAGATGGAGTCCAATGGAGGTTT---TG 1222
DB 1044 ATGGGATGGGCCAGCTGTCTACAGGCCAAGGATGGCATCTCAGTCAAGGGCGCAACA 1103
QY 1223 GATTTATAGTCCCAACCTCCAGCCCTCTTCTAAGGAGGAGAGAAACACCGGGAAC 1282
DB 1104 ACAGTAAACACCAACCCCTCTCCACCATCTAAGTCCCCCAGAGGAGATCGAAGAAC 1163
QY 1283 TCTACTGACATGAGCAAGAGAAATTGACACCATCTCCGGGCTGTTTCCCTTTCACTT 1342
DB 1164 TCTTATCCAGAGGCGCAAGAGATCGAATAATATCCGCATGGCTTCCCTGCTGCTCT 1223
QY 1343 TCCTCATCTTCAATATCTTCTACTGGGTTGTCTATAAAGTGTCTATGGTCAGAGATATCC 1402
DB 1224 TCCTCATTTCAACATGTTCTACTGATCATCTACAAGATTGTCCTGAGAGAGAGCTCC 1283
QY 1403 ACCA 1406
DB 1284 ACAA 1287

```

```

RESULT 3
AY418496
LOCUS
DEFINITION Mus musculus Glral gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY418496
VERSION AY418496.1 GI:39774456
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1294)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios

```

JOURNAL	Science 302 (5652), 1960-1963 (2003)
PUBMED	14671302
REFERENCE	2 (bases 1 to 1294)
AUTHORS	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferreria, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J., Adams, M.D. and Cargill, M.
TITLE	Direct Submission
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT	These sequences were made by sequencing genomic exons and ordering them based on alignment.
FEATURES	Location/Qualifiers
source	1..1294
gene	/organism="Mus musculus" /mol_type="genomic DNA" /db_xref="taxon:10090" <1..>1294 /gene="GLR1" /locus_tag="HMC6576"
ORIGIN	
Query Match	39.9%; Score 654.6; DB 9; Length 1294;
Best Local Similarity	72.8%; Pred. No. 2e-171;
Matches	941; Conservative 0; Mismatches 299; Indels 57; Gaps 5;
QY	116 CCAGGCCATGTCCTCTGATTTCTAGACAACTTATGGGGGAAACATCTGGATATG 175
DB	42 CCNAGCTATGTCACCTCTGGACTTCTGTGATGATCTATGGGAGGACTTCTGGGTATG 101
QY	176 ATGCCAGGATTCGGCCCAATTTTAAAGGCCACCCGTGAACGTGACCTGCAACATCTTCA 235
DB	102 AGCCAGGATCAGGCCCACTTTTAAAGGCTCTCTGTGAATGTAAGTTCGAACATCTTCA 161
QY	236 TCNACAGTTTCAGTCCGTGTCACCAAGACCAATGACCTACCGGTGAATGCTTCCTGC 295
DB	162 TCAACAGTTTCGGTTCATCGCTGAGACCAACCATGACTATAGGGTCAACATCTTCCCTGA 221
QY	296 GGCAACAGTGGATGATGACCAACGCTCTCTACCGAGAATATCTGATGACTCTCTGGACC 355
DB	222 GGCAGAGTGGAAAGACCCCGCTCTGCTTACATGATGATGATGATGATGATGATGATGATG 281
QY	356 TGATCCCTCCATGCTGGACTATCTGGAAGCCAGACCTCTTTTGTCTTAATGAGAAAG 415
DB	282 TTGACCCATCTATGTTGGATTCATCTGGAAGCTGACTGTTCTTTTGGCAATGAGAAG 341
QY	416 GGGCCAACTTCCATGAGTGGACCAACGACCAAGTACTGCGCATCTTCAGAATGGGA 475
DB	342 GGGCCCACTTCCAGAAATCACCGGACCAACAACTGCTTAAGAAATCTCCCGGAATGGCA 401
QY	476 ATGTGCTGTACAGCATCAGCTGACCTCATTTTGTCTGCTGATGGACCTCAAGAACT 535
DB	402 ATGTCTCTACAGCATCAGATCACCTGATCGTGGCTGCTGCCCATGGACCTGGAAGAT 461
QY	536 TCCCATGGACATCCAGAGTGCAGCATGCGAGCTTGAGAGCTCATCCATCTCTGCAGCC 595
DB	462 TCCCGATGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 503
QY	596 CTCTGCCATCTGTCTACCTTTCAGTTGGCTACACCATGAAAGACCTCTGTTGATGGCC 555
DB	504 TTTGGATATACCATGAAAGACCTCTATCTTTTGGTGGC 539
QY	656 TGGAGATGCTCTCTGTCCTGTCCTGAGGGGTGATCTGTCGCCCTGATCTGCTGCTGCTGCTG 715
DB	540 AGGAGCAG---GAGTGTGAGTGTGAGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 596
QY	716 GGGATGAGAAGGATCTAGGCTGTGTGTCACCAAGCACTACCAACAGGGAATTCACCTGCA 775
DB	597 AGGAAGAGAAAGACCTGAGATCTGACCAACCAAGCACTACCAACAGGTAATTCACCTGCA 656
QY	776 TCGAGTAAGATTTTACCTTGGAAAGCGCAGATGGGCTACTATCTGATTCAGATGTACATCC 835
DB	657 TCGAGGCCGATTTCCACCTGGAGCGCAGATGGGCTACTACCTGATCCAGATGTACATCC 716
QY	836 CCAGCCTACTCATCGTCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 895
DB	717 CCAGCCTGCTCATCGTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 776
QY	896 CTGCCCCGTGTGGGCTGGGCATCACACCGTGTCTACCATGACCAACCAAGAGCTCTGGCT 955
DB	777 CAGCTCGTGTGGGCTGGGCATCACACAGTGTCTACCATGACCAACCAAGAGCTCTGGCT 836
QY	956 CCGGGCCCTTTTGGCTAAGGTGTCTACGTGAAGGCAATCGACATCTCGATGGCTGTGT 1015
DB	837 CCGGAGCCTTCCCTTACCAAGGTGTCTACGTGAAGCTATTGACATTTGGATGGCTGTGT 896
QY	1016 GTCTGCTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1075
DB	897 GCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 956
QY	1076 AGCATTAAGAATTCATACGACTTCGAGAAGGAGGAGGCGCAACGCTTGGAGGAAGATA 1135
DB	957 AACACAAGGA-----ACTGCTTCGATTTAGGAGGAAGAGGACATCAAGGATGATG 1010
QY	1136 TCATCCAGAAAGTCTGTTTCTATTTCCGTTGGCTATGCTTGGGCTTGGGCT---ACTGCTGCAGG 1192
DB	1011 AGGTGGAGAGGCGCTTCAACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1070
QY	1193 CAAGAGATGGAGGTCCTCAATGGAAGTTCCTGGAATTTATAGTCTCC---CAACCTCCAGCCC 1249
DB	1071 CCNAGGATGCTCTCTGTCAGGGTGCACCAACCAACCAACCACTAACCCGCTCTG 1130
QY	1250 CTCCTTAAAGGAAGGAGAAACACGCGGAAACTCTACCTGACTGAGCAAGAGAAATG 1309
DB	1131 CGCCATCCAACTCCCGGAGGAGATCGGAAACTCTTTCATCCAGAGAGCAAGAGATCG 1190
QY	1310 ACACCATCTCCCGGCTGCTTCCCTTTCACCTTCTCATCTTCTCATATCTTCTACTGG 1369
DB	1191 ACAAGATATCTCGATCGGTTTCCCATGCTTCTCATCTTCTCATCTTCTCATCTTCTACTGGA 1250
QY	1370 TTGTCTATAAAGTGTATGCTCAGAGATATCCACCA 1406
DB	1251 TCATCTATAAGATCGTCCGAGAGAGGATGTCCACAA 1287
RESULT 4	
AK078258	2387 bp mRNA linear HTC 03-APR-2004
LOCUS	Mus musculus adult male olfactory brain cDNA, RIKEN full-length
DEFINITION	enriched library, clone:643054SK04 product:glycine receptor, alpha
ACCESSION	AK078258
VERSION	AK078258.1 GI:26097775
KEYWORDS	HTC; CAP trapper.
SOURCE	1 subunit, full insert sequence.
ORGANISM	Mus musculus (house mouse)
REFERENCE	1 Carninci, P. and Hayashizaki, Y. Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	1 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
JOURNAL	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
MEDLINE	Genome Res. 10 (10), 1617-1630 (2000)
PUBMED	20499374
REFERENCE	2 11042159
AUTHORS	3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,



QY 1250 CTCTTCTAAGGAGGAGAACACCGGAAACTCTACGTGGACTGACCAAGAGAAATTG 1309  
 Db 1650 CGCCATCAAGTCCCGGAGGAGATCGGAAACTCTTCCAGAGAGCAAGAGATCG 1709  
 QY 1310 ACACCATCTCCCGGCTGTCTCTCCCTTTCATCTTCTCTCAATCTCTTCTACTGGG 1369  
 Db 1710 ACAAGATATCTCGCATCGGTTTCCCATCGGCTTCTCTCTCTCAATCTCTTCTACTGGA 1769  
 QY 1370 TTGTCTATAAAGTGCTATGTCAGAGAGATATCCACCA 1406  
 Db 1770 TCATCTATAAGATCGTCCGAGAGAGATGTCCACAA 1806

RESULT 5  
 BC045547 3220 bp mRNA linear HTC 12-OCT-2004  
 LOCUS Homo sapiens cDNA clone IMAGE:5302075, containing frame-shift errors.

ACCESSION BC045547  
 VERSION BC045547.1 GI:28374243  
 KEYWORDS HTC.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 3220)  
 AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Udwin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raja, S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McSwan, P.J., McKernan, K.E., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickinson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E., Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.  
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002) 12477932

TITLE  
 JOURNAL  
 PUBLISHED  
 REFERENCE  
 DIRECTOR MGC Project.  
 DIRECT SUBMISSION  
 SUBMITTED (31-JAN-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK COMMENT  
 NIH-MGC project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
 cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN)  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Institute for Systems Biology  
<http://www.systemsbio.org>  
 Contact: [amadan@systemsbiology.org](mailto:amadan@systemsbiology.org)  
 Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodriguez, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAC Plate: 106 Row: p Column: 9  
 This clone was selected for full length sequencing because it

passed the following selection criteria: matched mRNA gi: 40254460  
 This clone has the following problem: frame shifted.

FEATURES  
 source  
 1. 3220  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5302075"  
 /tissue\_type="Brain, hypothalamus"  
 /clone\_lib="NIH MGC\_96"  
 /lab\_host="DH10B"  
 /note="Vector: pBluescript"

ORIGIN  
 Query Match 38.9%; Score 638.2; DB 3; Length 3220;  
 Best Local Similarity 70.0%; Pred. No. 1.1e+166;  
 Matches 942; Conservative 0; Mismatches 348; Indels 55; Gaps 4;  
 QY 63 CCTCCTCAGGGTGGCTTGGCAAAAGAGAAAGTCAAATCTGGAACCAAGGGGTCCAGCC 122  
 Db 600 CTTGAGGACGGCTTTCTGCAAAAGACCATGACTCCAGGTCTGGAAAACAACCTTTCAGAG 659  
 QY 123 CATGTCCTCTCTGATTTCTAGACAAACTTATGGGGCGAACTCTGGATATGATGCCAG 182  
 Db 660 CCTATCTCTCTCAGATTTCTTGGACAAGTTAATGGGAAGGACATCAGGATATGATCAAG 719  
 QY 183 GATTGGGCCCAATTTTAAAGGCCCAACCCCTGGAAGCTGACCTGCAACATCTTTCATCAACAG 242  
 Db 720 AATCAGGCCAAATTTTAAAGGCTCTCCAGTAAACGTTTCTTGCATATATTTTATCAACAG 779  
 QY 243 TTTTCAGTCCGTCCACCAAGACCAATGAGTACCGGGTGAATGTCTTCTTGGGCAACA 302  
 Db 780 TTTTGGATCAGTCACAGAAACGACCATGACTACCGAGTGAATATTTTCTGAGACAACA 839  
 QY 303 GTGGAATGACCAACGGCTGTCTTACCGAATATCTCTGATGATCTCTCGACCTCGATCC 362  
 Db 840 GTGGAATGATTCACGGCTGGCTGACAGTACCGAGTACCGAGTACCTCCCTGGACTTGGACCC 899  
 QY 363 CTCCATGCTGGACTCTATCTGGAAGCCAGACCTCTCTTTGCTTAATGAGAAAGGGGCCAA 422  
 Db 900 ATCCATGCTAGACTCCATTTGGAAACCATGTTGTTCTTTGGCAATGAGAAAGGGTCCAA 959  
 QY 423 CTTCCATGAGGTGACCAACGACCAACCAAGTTACTGGCATCTTCAAGAAATGGGAATGTCT 482  
 Db 960 CTTCCACGATGTCACCACTGACCAACAAATGCTGCGAATTCGAAAATGGCAAGTGTCT 1019  
 QY 483 GTACAGCATCAGGCTGACCCCTCATTTTGTCTGCTGATGGACCTCAAGAACTTCCCAT 542  
 Db 1020 CTACAGTATCAGACTCACCTTGCACCTTATCTCTGCGATGGACTTGAAGAACTTCCGAT 1079  
 QY 543 GGACATCCAGACGTGCAGCATGCAGCTTCAGAGCTCATCCATCTCTGAGCCCTCTGTC 602  
 Db 1080 GGATGTCAGACTCTGTACAAATGAGCTGGAGATT----- 1114  
 QY 603 ATCTCTGTCACTTTTCAGTTGGCTTACACCATGAAAGACCTCGTGTGTTGAGTGGCTGGAAGA 662  
 Db 1115 -----TTGGGTACAGCATGAATGACCTGTATTTGAGTGGTAAAGTGA 1157  
 QY 663 TGCTCTGCTGTCCAAGTGGCTGAGGGGCTGACTCTGCCCCAGTTTATCTTGGGGATGA 722  
 Db 1158 TGGTC---CAGTGCAAGTTGCTGGAAGGATGACCCCTGCCCCAGTTTATTTTGAAGAAGA 1214  
 QY 723 GAAGGATCTAGGCTGTGTACCAAGCACTACAACACAGGGAATTCACCTGCAATCGAGGT 782  
 Db 1215 GAAGGACTGGCTACTGTACAAAGCACTACAACACTGGAAGATTTTACCTGCAATCGAGGT 1274  
 QY 783 AAAGTTTCACTGGAAACGGCAGATGGGCTACTATCTGATTCAGATTAATATCCCAAGCCT 842  
 Db 1275 CAAAGTTTCACTGGAAACGGCAAAATGGGATATTTATTTGATCCAGATGTACAT-CCAGCCT 1333  
 QY 843 ACTCATGCTCATCTGCTCTGGGCTCTCTTCTGATCAACATGATGCTGCCCTCCCG 902  
 Db 1334 GCTTATAGTAATTTTGTCTCTGGGTTTCTTTTGGATAAATATGATGTCAGCCCTCCAG 1393



```
QY 903 TGTGGCGCTGGGCGATCACACCGTGTCTACCATGACCAACCGAGCTCTGGCTCCCGGGC 962
Db 1394 GGTGCGCATGGGCGATCACACAGTCTTAACGATGACCAACCGAGTTCAGGCTCCAGGGC 1453
QY 963 CTCCTTTGCTAGGCTGCTTACGTGAAGCAATCGACATCTGATGGCTGTGTCTGTCT 1022
Db 1454 ATCTCTGCAAGGCTCTCTATGTAAGAGGATGACATCTGATGGCGGTGTGCTCTCT 1513
QY 1023 CTTTGTGTTGCTGCTTCTGCGAGTATGCTGCATATAAATTTTGTCTCGTCAGCATAA 1082
Db 1514 GTTTGTGTTGCTGCTTACTGGAATACGAGCGGTGAATCTGCTCTCCAGGCAACACAA 1573
QY 1083 AGAATTCATACGATCTCGAAGAGGAGGCGCCCAACGCTTGGAGGAAGATATCATCCA 1142
Db 1574 GGAGTTCTCTGCGCTCCGAGAGAGACAGAGAGGAGCAATAAGAGAGAGACGTTTACTCG 1633
QY 1143 AGAAGCTGCTTTCTATTTCTGCTGCTATGCTTGGCGCACTGCTTCCAGGCAAGATG 1202
Db 1634 TGAAGTCTGTTTAAATTTTACGCTTATGGGATGGGTCACTGCTCCAAAGTGAAGATGG 1693
QY 1203 AGTCCAAATGGAGGTTCTGGCATTTATATGCTCCCAACCTCCAGCGCCCTCTTCTAAGGA 1262
Db 1694 AACAGCTGTCAAGGCCAC-----ACCTGCCAACCCCACTCCCAACCGCGCAAGA 1744
QY 1263 AGAGAAACACCGCGAACTCTACGTGACTGAGCCAGAGAAATTCACACCATCTCCCG 1322
Db 1745 TGAGATGCTATCAAGAAGAAGTTTGTGGACCGGCAAAAGGATTCACACGATATCTCG 1804
QY 1323 GGTGCTTCTTCCCTTCACTTCTCTCATCTTCAATATCTTCTACTGCTGTCTATAAGT 1382
Db 1805 AGTGGCTTCCCATGCGCTTCTCTCATTTTCAACATCTTTTACTGGATCACATACAGAT 1864
QY 1383 GCTATGCTCAGAGATATCCACAG 1407
Db 1865 CATTCGGCATGAGATGTCACAG 1889

RESULT 6
AY418495 1294 bp DNA linear GSS 17-DEC-2003
LOCUS Pan troglodytes GLRA1 gene, VIRUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY418495
VERSION AY418495.1 GI:39774455
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferred nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
2 (bases 1 to 1294)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
These sequences were made by sequencing genomic exons and ordering
them based on alignment.
COMMENT
FEATURES
source
Location/Qualifiers
1..1294
/organism="Pan troglodytes"
/mol_type="genomic DNA"
```

```

/db_xref="taxon:9598"
<1..>1294
/genes="GLRA1"
/locus_tag="HCM6576"

gene
ORIGIN
Query Match 38.5%; Score 630.8; DB 9; Length 1294;
Best Local Similarity 67.2%; Pred. No. 9.4e-165;
Matches 886; Conservative 0; Mismatches 387; Indels 45; Gaps 2;

QY 89 AGGAAGTCAAAATCTGGAACCAAGGGGTCCAGCCCATGTCCCTCTGATTTCTTAGACA 148
Db 15 AGAGGCTGAAGTGTCTCGCTCCGACCAAGCCCTATGTACCCCTCGGATTTCTTGATA 74
QY 149 AACTTATGGGCGAATCTGATATGATGCGAGGATTCGCGCCAAATTTTAAAGGCCAC 208
Db 75 AGCTAATGGGAGAACCTCCGATATGATGCGAGGATTCAGGCCCAATTTNNNNNGTCCC 134
QY 209 CCGTGAACGTGACCTGCAACATCTTCATCAACAGTTTCAGCTCCGTCAACAGACCACAA 268
Db 135 CAGTGAACGTGAGCTGCAACATTTTCATCAACAGCTTTGGTTCATTTGTGAGACAACA 194
QY 269 TGGACTTACCGGCTGAATGTCTTCTGCGGCAACAGTGGGAATGACCCACGCTCTCTTACC 328
Db 195 TGNNNNNNGSGTCAACNNCTTCTGCGGCGCAATGGAACGACCCCGCTGCGCTATA 254
QY 329 GAGAATATCTGATGATCTCTGGACTCGATCCCTCCATGCTGGAATCTATCTGGAAGC 388
Db 255 ATGAATACCTGATGATCTCTGGACTCGATCCCTCCATGCTGGAATCTATCTGGAAGC 314
QY 389 CAGACTCTCTTCTGATGATGAGGAGGCGGCAACTTCATGAGGTGACCAACGAGCAACA 448
Db 315 CTGACCTGCTTCTGCAACAGAGAGGCGGCGCACTTCCATGAGATCAACACAGACAACA 374
QY 449 AGTTACTGCGCATCTTCAAGAAATGGGAATGTGTGTACAGCATCAGCTGACCTCATTT 508
Db 375 AATTGCTAAGATCTCCCGAATGGGAATGTCTCTACAGCATCAGATCACCCTGACAC 434
QY 509 TGTCTGCTGATGAGCACTCAAGAACTTCCCATGAGATCCAGAGCGTGACGATGACGAGC 568
Db 435 TGGCTGCGCCATGGACTTGAAGAAATTTCCCATGATGTCCAGACATGTATCATGCAAC 494
QY 569 TTGAGAGCTATCCATCTCTGCGAGCCCTCTGCCATCTCTGTACATTTTCACTTGGCTACA 628
Db 495 TGGAAAGCT-----TTGGATATA 512
QY 629 CCATGAAGACCTCGTGTGAGTGGCTGGAAGATGCTCTCTGCTGCTCAAGTGGCTGAGG 688
Db 513 CGATGAATGACCTCATCTTTGAGTGGCAGGAACAGG---GAGCGGTGACAGTACAGATG 569
QY 689 GGCTGATCTGCGCCAGTTTATCTTTCGGGATGAGAAAGATCTAGGCTGTGTGACCAAGC 748
Db 570 GACTAACNNNGCCCGAGTTTANNNNNAAGAGAGAGAGGACTTTGAGATNNNTGACCAAGC 629
QY 749 ACTAGAACACAGGGAATTCACCTGATGAGTAAAGTTTACCTTGGAGCGGAGATGG 808
Db 630 ACTACAACACAGGTAATTCACCTGATGAGGCGCGGTTCCACCTTGGAGCGGAGATGG 689
QY 809 GCTACTATCTGATTCAGATGTATATCCCGAGCTCTCATCTGCTCATCTCTCTGGGTCT 868
Db 690 GTTACTACTGATTCAGATGTATATTCACAGCTGCTCATTTGTCATCTCTCATGATCT 749
QY 869 CTTCTGATCAACATGATGCTGCGCCCTGCGCGCTCTTTGCTTGGAGTGTCTTACGTA 928
Db 750 CTTCTGATCAACATGATGCTGCGCCCTGCGCGCTCTTTGCTTGGAGTGTCTTACGTA 809
QY 929 TCACCATGACACAGCTGCTGCGCGCTCTTTGCTTGGAGTGTCTTACGTA 988
Db 810 TCACCATGACACAGCTGCTGCGCGCTCTTTGCTTGGAGTGTCTTACGTA 869
QY 989 AGGCAATGACATCTGATGCTGCTCTTTGCTTGGAGTGTCTTACGTA 1048
Db 870 AGGCAATGACATCTGATGCTGCTCTTTGCTTGGAGTGTCTTACGTA 929
```





483	Db	CTACGATATCAGACTCACCTTGACCTTATCTGTGCCATGGAGCTTGGAAGAACTTTCCGAT	542
543	Qy	GGACATCCAGAGCTGCAGATGCAGCTTGAGAGCTCATCCATACTCGAGCCCTCTGCC	602
543	Db		
543	Db	GGATGTCAGAGCTGTACAATGCAGCTGGAGATT	577
603	Qy	ATCTCTGTCACTTTAGTTGGCTACACCAATGAAGAAGCTCGTGTTTGAAGTGGCTGGAAGA	662
578	Db		
578	Db	-----TTGGGTACACGATGAATGACCTGATATTTTGAAGTGGTAAAGTGA	620
663	Qy	TGCTCTCTGTGTCACCAAGTGCCTGAGGGCTGACTCTGCCCCAGTTTATCTTTCGGGGATGA	722
621	Db		
621	Db	TGATC---CAGTGCAAGTTGCTGAAGATGACCTTGCCCCAGTTTATTTTGAAGAAGAGA	677
723	Qy	GAGGATCTAGGCTGTGTATCCAAAGCACTACAAACACAGGGAATTCACCTGCATCGAGGT	782
678	Db	GAAGGAACCTGGCTACTGTACAAAGCACTACAACACTGGAAAGTTTACCTGCAATTCAGGT	737
783	Qy	AAAGTTTCACTGGAAACGGCAGATGGGCTACTATCTGATTCAGATGTACATCCCCAGCCT	842
738	Db		
738	Db	CAAGTTTTCATCTGGAAACGGCAATGGGATATTATTTGATCCAGATGTACATCCAGCCT	797
843	Qy	ACTCATCGTCATCTCTGCTGGTCTCTCTTCTGGATCAACATGGATGCTGCCCTGCCCG	902
798	Db		
798	Db	GCTTATAGTAATTTTGTCTGGTTCCTTTTGGATTAATATGATGGATGCCCTGCCAG	857
903	Qy	TGTGGGCTCGGGCATCACACCGTGTCTACCATGACACCCAGAGAGCTCTGGCTCCGGGC	962
858	Db	GGTCGCACTGGGCATCACACAGTCTTAACGATGACACCCAGAGTTTCAGGCTCCAGGGC	917
963	Qy	CTCTTTGCTTAGGTTGCTCTACGTGAAGGAATCGACATCTGGATGGCTGTGTGTCTGCT	1022
918	Db		
918	Db	ATCTCTGCCAAAGTCTCCTATGTATAAAGCGATTGACATCTGGATGGCGGTGTGCTCTCT	977
1023	Qy	CTTTGTGTTGCTGCTCTCTCGAGTATGCTGCCATAAAATTTGTTTCTCGTCAGCATAA	1082
978	Db		
978	Db	GTTTGTGTTTGTGCTTCTACTGATACGAGCGGTGAATCTGCTCCAGGCAACACAA	1037
1083	Qy	AGAATTATACAGACTTCGAAGAAGCGAGCGGCCA	1118
1038	Db	GGAGTTCTCGGCTCCGAGAAGACAGAAGGCCA	1073
RESULT 9			
AY419225			
LOCUS			
DEFINITION	AY419225 1398 bp DNA linear GSS 17-DEC-2003		
	Mus musculus GUR3 gene, VIRTUAL TRANSCRIPT, partial sequence,		
	genomic survey sequence.		
ACCESSION	AY419225		
VERSION	AY419225.1 GI:39775185		
KEYWORDS	GSS.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	1 (bases 1 to 1398)		
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Perrier,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.		
	Inferring nonneutral evolution from human-chimp-mouse orthologous		
	gene trios		
TITLE	Science 302 (5652), 1960-1963 (2003)		
JOURNAL	14671302		
PUBMED	2 (bases 1 to 1398)		
REFERENCE	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Perrier,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.		
AUTHORS	Direct Submission		
TITLE	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,		
JOURNAL	Rockville, MD 20850, USA		
	These sequences were made by sequencing genomic exons and ordering		
COMMENT			



	992	CACCTTCTGAGATATGCAGCTGTAAATTTTGTAATCAAGAACAACAAGAACCTTCCTGAGAT	105
Df	1097	TTTGAAGAAGCAGAGGGCCCA-----ACGCTTGGAGGAAGATATCATCC	1141
Qy	1052	TTTGGCGAAGAAGAAGAATAAGACAGAAGCTTTTGGACTGGAGAAGTTTTACCCTTCT	1111
Df	1142	AGAAAAGTCGTTTCTATTCTTCGCTGGCTATGGCTGGGCCATCTGCTGCGCAAGAGATG	1201
Qy	1112	CAGATATGATGATGAGGTAAAGGAAAGCCGATTCAGCTTTCACAGCCTATGGAATGGAC	1171
Df	1202	GAGTCCCATGCAAGGTTCTGCGCAT-----TTATAGTCCCCCAACTCCAGCCC--TC	1252
Qy	1172	CATGCTTACAGCAAAGATGGCATGATCCAAAGGGGCCCAACACCTCTCCAGGTAA	1231
Df	1253	TTCTAAGGGAAGGAGAAACCAACGCGAAACTCTACGTGGAATGAGCCAAGAGAAATTGACA	1312
Qy	1232	TGCCAAAAGTCTGATGNAATGAGGAAGTCTTTATCGACGGGCCCAAGAGATTGATA	1291
Df	1313	CCATCTCCCGGCTGTCTTCCCTTTTCACTTTTCTCTCATCTTCAATATCTTCTACCTGGTGTG	1372
Qy	1292	CCATCTCCCGGCTGTCTTCCCTTTTCACTTTTCTCTCATCTTCAATATCTTCTACCTGGTGTG	1351
Df	1373	TCTATAAGTGTATGTCAGAGATATCCACCAG	1407
Qy	1352	TCTATAAATTTCTTAGGCATGAGGATATTCATCAG	1386
Df			
RESULT 11			
LOCUS	AY419224		
DEFINITION	Pan troglodytes GLRA3 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.		
ACCESSION	AY419224		
VERSION	AY419224.1	GI:39775184	
KEYWORDS	GSS.		
SOURCE	Pan troglodytes (chimpanzee)		
ORGANISM	Pan troglodytes		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan. 1 (bases 1 to 1382) Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarialwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios		
AUTHORS	Science 302 (5652), 1960-1963 (2003) 14671302		
JOURNAL	2 (bases 1 to 1382)		
PURNED	Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarialwal,A., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. Direct Submission		
REFERENCE	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA		
AUTHORS	These sequences were made by sequencing genomic exons and ordering them based on alignment.		
TITLE	Location/Qualifiers		
JOURNAL	source		
COMMENT	/organism="Pan troglodytes" /mol_type="genomic DNA" /db_xref="taxon:9598"		
FEATURES	<1..>1382 /gene="GLRA3" /locus_tag="HCM6819"		
ORIGIN			
Query Match	28.0%;	Score 458.8;	DB 9; Length 1382;
Best Local Similarity	61.0%;	Pred. No. 1.4e-116;	
Matches	824; Conservative	0; Mismatches 455;	Indels 72; Gaps 7;
QY	80	TGCGAAAGAGGAGTCAAACTCTGGAACCAAGGGTCCAGCCCATGTCCTCTGATT	139
Df	77	TTTCCCAAAAGGAAAAAGACAGTGCAGAGTCTCGAAGTGTCTCAATGTCACCTTCTGATT	136
QY	140	TCTTACGACAACTTATGGGGCGAATCTGGATATGATGCCAGGATTCGGGCCCAATTTTA	199
Df	137	TTCTGGATAAATTAATGGCGGAGACATCAGGATATGATGCAAGATCAGACCCCAATTTTA	196
QY	200	AGGCCCAACCGGTGAAGTGCCTGCAACATCTTCATCAACAGTTTTCAGCTCCGTCAACA	259
Df	197	AAGNN	256
QY	260	AGACCAATGGAATCGGGTGAATGTCTTTCGGGCAACAGTGGGAATGACCCAGCC	319
Df	257	NNNNNNNNNNGATTACAGAGTGAATATCTTCTTCGTGAGAAATGGAATGATCCCGCC	316
QY	320	TGTCCTACGAGAAATCTCTGATGATCTCTGGAACCTGATCCCTCCATGCTGACATCTA	379
Df	317	TCGCGTACAGTGAATATCTGACGACTCTTTAGACCTCGACCCCTCCATGTTGGACTCCA	376
QY	380	TCGGAAGCAGACCTCTTCTTGTCTTAATGAGAAAGGGGCCAACTTCCATGAGTGACCA	439
Df	377	TTTGGAAACCTGATTTGTTCTTTGGCAATGAAAGGGTCCCACTTTCATGAGTCACTA	436
QY	440	CGGACAAAGTACTGCGCATCTTCAAGAAATGGAATGTGCTGACAGCATCAGG---C	496
Df	437	CAGACAAATTTGCTAAGATTTTCAAAATGGAATGTTCTTTATTCATATAGGAGAT	496
QY	497	TGACCTCATTTTGTCTGCTGATGAGACCTCAAGAACTTCCCAATGAGCATCACAGCT	556
Df	497	TAAATTAACACTTCTCTGCAATGATGATCTCAAGAAATTTTCCCATGATGATCAAAACAT	556
QY	557	GCAAGTGCAGCTTGAGAGCTCATCATCTCTGAGCCCTCTGCCACTCTCTGTCACTTT	616
Df	557	GTATAATGCACTGGAAAGCT-----	577
QY	617	CAGTTGGCTACACATGAAAGACCTCGTGTGAGTGGCTGGAAGATGCTCTGCTGTCC	676
Df	578	---TTGGGTACAAATGATGATCTCATTTTGAATGGC---AAGATGAGGACCCGCTAC	631
QY	677	AAGTGGCTGAGGGCTGATCTGTCGCCAGTTTATCTTGGGGATGAGAAGATCATAGCT	736
Df	632	AAGTGGCAGAAGGACTCACTTTGCCCGCAGTTTCTGTGAAAGAAAGAAAGATTAAGAT	691
QY	737	GTGTACCAAGCACTACACAGAGGAATTCACCTGATCGAGTAAGTTTCACTGG	796
Df	692	ACTGCACTAAACATTAACATAAGGAAATTTACGTGTATAGAGTGCATTCCTG	751
QY	797	AAGCGCAGATGGGCTACTATCTGATTACAGATGATATCCCAAGCCTCTCATCTGTCATCC	856
Df	752	AGCGCAATAGGATATCTGATCCAGATGATCAATTCCTGATCTCTGATGTTATTC	811
QY	857	TGTCCTGGGTCTCTTCTGATCAACATGATGCTGCCCTGCGGTGGGCTGGGCA	916
Df	812	TATCTCTGGTTCGTTCTCGATCAACATGATGACAGCCGCGGAGGTAGCTCTGGGGA	871
QY	917	TACACCGTGTCTACCATGACCAAGGAGCTCTGGCTCCCGGCGCTCTTGGCTAAGG	976
Df	872	TAAACCAACCGTGTAAAGATGACTTACACAGAGTTTACAGGATCAAGAGTCTTTCGCAAAAG	931
QY	977	TGTCCTACGTGAGGCAATCGACATCTGATGCTGTGTCTGCTCTTCTGTTTCTGCTG	1036
Df	932	TTTCATATGCAAGCTATTTGATATTTGGATGCGATATGCTCTCTCTTGTGTTTCAG	991
QY	1037	CCTTGTCTGAGTATGCTGCCATAAAATTTTGTCTCTGTCAGCATAAAGATTCATACGAC	1096

Db 77 TTGCCAAGGAAACAGATAGTGAAGATCTGAAAGTGTCTCAATGTCACTTCTGATT 136  
Qy 140 TCTAGACAACTTATGGGCGGAACATCTGGATATGATGCCAGGATTCGGCCCAATTTTA 199  
Db 137 TTTGATGAAATTAATGGGCGAGACATCAGGATATGATCAAGAAATCAGACCCCAATTTTA 196  
Qy 200 AAGGCCACCCGTGAACCTGACCTGCAACATCTTCATCAACAGTTCAGCTCCGTCAACA 259  
Db 197 AAGNN 256  
Qy 260 AGACCAATAGGACTACCGGGTGAATGTCTTTCTTGCGGCAACAGTGGAAATGACCCAGCC 319  
Db 257 NNN 316  
Qy 320 TGTCTCTACGAGATATCTGTGATCTCTGTGACCTCGATCCCTCCATCTGCTGACTCTA 379  
Db 317 TCGCGTACAGTGAATATCTGTGAGCTCTTTAGACCTCGACCCCTCCATGTTGGACTCCA 376  
Qy 380 TCTGGAAGCCAGACCTCTTCTTGTCTAATGAGAAAGGGCCAACTTCCATGAGGTGACCA 439  
Db 377 TTTGGAACCTGATTTGTTCTTCCCAATGAAAGGGTGCACACTTTTCATGAGTCACTA 436  
Qy 440 CGGACAAAGTTPACTCGGCATCTTCAAGAAATGGGAATGTCTGTGACAGCATCAGG- -C 496  
Db 437 CAGACAACAAAATGCTAAGAAATTTCAAAAAATGGAATGTCTTTATTCAATAAGGAGAT 496  
Qy 497 TGACCCCTCATTTGTTCTGCTGATGACCTCAAGAACTTCCCATGGACATCCAGACGT 556  
Db 497 TAACTTAAACACTTCTCTGCAATGGATCTCAAGAAATTTCCCATGGATGACAAACAT 556  
Qy 557 GCACGATGAGCTGTGAGAGCTCATCATCTCTGCAGCCCTCTGCCATCTCTGTCACTTT 616  
Db 557 GTATATGCACTGGAAAGCT----- 577  
Qy 617 CAGTTGGCTACACCATGAAGACCTCGTGTGAGTGGCTGGAAGATGCTCCTGTGCTC 676  
Db 578 ---TTGGGTACACAATGAATGATCTCATTTTGAATGGC---AAGATGAGGACCCCGTAC 631  
Qy 677 AAGTGTGAGGGGCTGACTCTGCCCACTTATCTTGGGGATGAGAAGGATCTAGGCT 736  
Db 632 AAGTGTGAGGAGGACTCACTTTGCCCACTTCTGTGAAAGGAAAGAAATTTACGAT 691  
Qy 737 GTGTGTACCAAGACTCAACACAGGAAATTCACCTGCATGAGGTAAAGTTTCACTGTG 796  
Db 692 ACTGCACTAAACATTACATACAGGAAGTTTACGTGTATAGAGTGGCATTCATCTGG 751  
Qy 797 AACGGCAGATGGGCTACTATCTGATTCAGATGTACATCCCGACCTTACTCATGTCATCC 856  
Db 752 AGCGACAATGGGATACTATCTGATCCAGATGTACATTCACAGTCTCTCTGATTGTTATTC 811  
Qy 857 TGTCTCTGGGCTCTCTTCTGATCAACATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 916  
Db 812 TATCTCTGGGTTCTGTTCTGATCAACATGATGATGATGATGATGATGATGATGATGATG 871  
Qy 917 TCACCAACCGTGTCTACATGACCAACAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 976  
Db 872 TAACCACTGTGCTAAAGATGACTACACAGATTCAGATTCAGATTCAGATTCAGATTCAG 931  
Qy 977 TGTCTCTAGTGAAGGCAATCGACATCTGAGTGGCTGTGTCTGTCTGCTGCTGCTGCTGCTG 1036  
Db 932 TTTTCATATGTCAAAGCTATTGATTTGGATGGCAGTATGCTCTCTCTCTCTCTCTCTCT 991  
Qy 1037 CTTGCTGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1096  
Db 992 CACTTCTGAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1051  
Qy 1097 TTGGAAGAAGG-----CAGAGGCGCAACCTGAGGAAAGATAT---CATCC 1141  
Db 1052 TTGCGAAGAGAGAAAGAAATAGACAGAGCTTTTGCATTTGGAGAGTTTACCGTTTCT 1111  
Qy 1142 AAGAAATCGTTTCTATTTCCGTGGCTATGGCTTGGGCCACTGCTGCGAGGCAAGAGATG 1201

Db 1112 CAGATATGATGATGAGGTAAAGGAAAGCCGATTTCAGCTTTCAGCTTATGGAATGGGAC 1171  
Qy 1202 GAGGTCCATGGAAGTCTCTGGCAT-----TTATAGTCCCAACCTCCAGCCCC---TC 1252  
Db 1172 CATGTCTACAGCAAGAGTGGCATGACTCCAAAGGCGCCCAACCACTGTTCAGGTAA 1231  
Qy 1253 TTCTAAGGGAAGGAGAAACACCGGAAACTCTAAGTGGACTGAGCAAGAGAAATTTGACA 1312  
Db 1232 TGCCAAAAGTCTCTGATGAATGAGGAAGTCTTTATCGACCGGCCANNNNNNNGACA 1291  
Qy 1313 CATCTCCCGGGCTGTCTCCCTTTTCACTTTCTCATCTTCAATATCTTCTACTGGGTG 1372  
Db 1292 CCATCTCCCGAGCTGCTTCCCATTTGATTTTGAATTTTAATATTTTCTACTGGGTTA 1351  
Qy 1373 TCTATAAGTGTCTATGGTCAAGAGATATCCA 1403  
Db 1352 TCTATAAAATTTCTAGGCATGANNATATCA 1382

## RESULT 12

CD354404  
LOCUS CD354404  
DEFINITION UI-M-GMO-cgd-k-09-0-UI.r1 NIH\_BMAP\_GMO\_Mus musculus cDNA clone  
IMAGE:30361304 5', mRNA sequence.  
ACCESSION CD354404  
VERSION CD354404.1 GI:31146905  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
REFERENCE 1. (bases 1 to 773)  
AUTHORS NIH-MGC <http://mgi.nhl.nih.gov/>  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-rc@mail.nih.gov](mailto:cgabbs-rc@mail.nih.gov)  
Tissue Procurement: Dr. Jim Lin, University of Iowa  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Distribution information can be found at  
<http://genome.uiowa.edu/distribution/mousef1.html>  
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: pYX-5.

FEATURES  
source

Location/Qualifiers  
1. .773  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6"  
/db\_xref="taxon:10090"  
/clone="IMAGE:30361304"  
/tissue\_type="whole brain"  
/dev\_stage="1, 5 and 15 days newborn"  
/lab\_host="DH10B (T1 phage resistant)"  
/clone\_lib="NIH\_BMAP\_GMO"  
/note="Organ: Brain; Vector: pYX-Asc; Site:1: Ecor I;  
Site 2: Not I; The library was constructed according  
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
1996. Denatured RNA was size fractionated on a 1% agarose  
gel. First strand cDNA synthesis was primed with oligo-dT  
primer containing a Not I site. Double strand cDNA was size  
selected according to mRNA size fraction, ligated with Ecor  
I adaptor, digested with NotI and then cloned  
directionally into pYX-Asc vector. The library tag  
sequence located between the Not I site and the polyA tail  
is CCAACTGAAT. This library was created for the University  
of Iowa Brain Anatomy Project (BMAP). 'Gene Discovery in the  
Developing Mouse Nervous System', supported by National  
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,  
program coordinator."

## ORIGIN

```

Query Match      27.0%; Score 442.4; DB 6; Length 773;
Best Local Similarity 74.2%; Pred. No. 4.5e-112;
Matches 607; Conservative 0; Mismatches 166; Indels 45; Gaps 2;

298 CAACAGTGAATACCCAGCCTGTCTACCGAGATATCTCTGATGACTCTCTGGACCTC 357
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   1 CAGCAGTGGATGATTCACGGCTGGATACAGTGTACCCAGATGATTCCTGGATTG 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

358 GATCCCTCAGTGTGACTCTATCTGGAAGCCAGACTCTCTTTTCTAATGAAAGGG 417
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   61 GATCCCTCAATGTGGATTGATTCGATTGAAACCGGATTGTTCTTTGCCAATGAAAGGGT 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

418 GCCAATCTCCATGAGTGTACCCAGGACCAACAGTTACTGGCATCTTCAAGATGGGAAT 477
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   121 GCCAATTCATGATGTACACCTGACCAACAGTTGTTGGGATTTCCAAAATGCGAA 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

478 GTCTCTGACAGCATCAGGCTGACCTCATTTGCTCTGCTGCTGATGACCTCAAGAACTTC 537
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   181 GTCTCTACAGTATTAGACTCACCCTGACTTATCTGTCCAATGACCTGAAGAACTTT 240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

538 CCATGGACATCCAGACGTGSCAGATGACAGCTTGAGAGCTCATCCATCTCTGACGCCCT 597
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   241 CCAATGGATGTCAGACCTGTACAATGACGTGGAGTT----- 280
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

598 CTGCCATCTGTCACTTTTCAGTTGGCTACACCATGAAGACCTCGTGTGAGTGGCTG 657
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   281 -----TTGGGTACACCATGAATGACCTGATATTGAGTGGTTA 318
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

658 GAAGATGTCCTCTGTCCTCAAGTGGCTGAGGGCTGACTCTGCCAGTTTATCTTGGCG 717
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   319 AGTATGGTC---CAGTACAAGTTGCTGAAGGACTCACCTTGCCCGAGTTTATTTGAAA 375
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

718 GATGAGAAGGATCTAGGCTGTTGTACCAAGCACTACAACACAGGGAATTCACCTGCATC 777
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   376 GAAGAGAGGAGCTTGGTTATTGCACAAAGCATTACACACTGGCAAGTTTACCTGCATT 435
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

778 GAGGTAAGTTTCACTGGAAACGGCAGATGGGCTACTATCTGATTCAGATGTAATCCCC 837
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   436 GAGGTCAAAGTTTCACTGGAGCCAGATGGGCTACTATTTGATCCAGATGTACATCCCC 495
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

838 AGCTTACTATCTGATCTGCTCTGCTGCTGCTCTCTGATCAACATGATGCTGCCCT 897
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   496 AGCTTGTGATGATCAATTTGCTGCTGCTGCTCTTTTGGATAAACAATGATGAGCCCT 555
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

898 GCCCGTGTGGCCCTGGGCATCACCAAGCTGCTCACCATGACCAACCCAGAGCTCTGGCTCC 957
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   556 GCCAGGTTGCCCTGGGCATCAACAGTCTGACATGACTACACAGATTCAAGTTCC 615
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

958 CGGGCTCTTTTGCTTAAGGTGCTCTACGTGAAGGCAATCGAATCTGGATGGCTGTGT 1017
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   616 AGGGCATCTCTGCCAAAGGCTCTCTATGTGAAGCAATTGACATCTGGATGGCGGTATGC 675
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

1018 CTGCTCTTTGTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1077
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   676 CTCTCTTTGTGTTGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 735
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

1078 CATAAAGAAATTCATAGACTTTCGAAGAGCAGAGCG 1115
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   736 CATAAGGAGTTCTTGGTCTCCGGAGACGACAGAGAG 773
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 13
AUI69868
LOCUS AUI69868 662 bp mRNA linear EST 29-JAN-2001
DEFINITION AUI69868 Ol-br-ad cDNA Oryzias latipes cDNA clone br5332, mRNA
sequence.
ACCESSION AUI69868
VERSION AUI69868.1 GI:12591937
KEYWORDS EST.
SOURCE Oryzias latipes (Japanese medaka)
ORGANISM Oryzias latipes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;  
 Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.

# REFERENCE AUTHORS TITLE JOURNAL COMMENT

Mita, K., Ishikawa, Y. and Yamauchi, M.  
 Establishment of cDNA database of medaka, *Oryzias latipes*  
 Unpublished (2001)  
 Contact: Mita K

Genome Research Group  
 National Institute of Agrobiological Sciences  
 Owashi 1-2, Tsukuba, Ibaraki 305-8634, Japan  
 Email: kmica@nias.affrc.go.jp  
 method: uni-directional sequence direction: sequenced from T3 primer  
 (5' -> 3')

## FEATURES source

Location/Qualifiers  
 1..662  
 /organism="Oryzias latipes"  
 /mol\_type="mRNA"  
 /strain="HNT"  
 /db\_xref="taxon:8090"  
 /clone="br5332"  
 /sex="female/male mixed"  
 /tissue type="brain"  
 /dev stage="adult"  
 /clone\_lib="Ol-br-ad cDNA"

## ORIGIN

```

Query Match      24.5%; Score 401.8; DB 1; Length 662;
Best Local Similarity 76.4%; Pred. No. 9.9e-101;
Matches 540; Conservative 0; Mismatches 122; Indels 45; Gaps 2;

195 TTTTAAAGCCCAACCCGTGAACGTGACCTGCAACATCTTCATCAACAGATTTCAGCTCCGT 254
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   1 TTTTAAAGTCCACCTGTAAAGCTCACCTGCAACATTTTATCAACAGCTTTGCGATCCAT 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

255 CACCAAGACCAATGAGCTACCGGTGAATGCTTTCTTGGCGCAACAGTGGAAATGACCC 314
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   61 CGCTGAACCAACATGGACTACAGATGAACATCTTCTGAGGACGACAGTGGAAACGACCC 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

315 AGCCCTGTCTTACCGAGAATATCTGATGACTCTCTGGACCTCGATCCCTCCATGCTGGA 374
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   121 CGGCTGGCTTACAGCGAGTATCCCGACGACTCGCTGGATCTGGACCCGTCATGTTGA 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

375 CTCATCTGGAAGCCAGACCTCTCTTTTGTGTAATGAGAAAGGGGCCAACATTCATGAGGT 434
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   181 CTCATCTGGAAGCCGATCTGTTCTTGTATGAGAAAGGGGCCAACATTCACAGAGGT 240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

435 GACCAAGGACCAACAGTACTCGGCATCTTCAAGATGGGAATGTCTGTACAGCATCAG 494
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   241 CACCAAGGACCAACAGCTGCGCATCTCCAAAATGGCAACGTGTGTACAGCATACG 300
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

495 GGTGACCTCATTTTCTGCTGCTGATGGACCTCAAGAACTTCCCATGGACATCCAGAC 554
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   301 AATCACTCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

555 GTGCAAGATGCAAGTGTGAGAGCTATCCATATCTCTGACGCCCTCTGCGCATCTCTGTCACT 614
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   361 CTGCATCATGCACTGAGAGCT----- 383
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

615 TTCAGTTGGCTACACCATGAAGACCTCGTGTGTTGAGTGGCTGGAGATGCTCTCTGCTGT 674
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   384 -----TCGGCTACACCATGAACGACCTCATCTCTCGAGTG---GGACGAGAAAGGGCCGCT 435
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

675 CCAAGTGGCTGAGGGGCTGACTCTGCCCGCATTTATCTTTCGGGGATGAGAAAGGATCTAGG 734
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   436 GCAGTGGCGGCGGCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 495
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

735 CTGTTGTACCAAGCACTACAACACAGGGAATTCACCTGCATCGAGGTAAAGTTTCACCT 794
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   496 CTACTGCAACCAAGCACTACAACACAGGTAATTCACCTGCTGATCGAGGCTCGCTTCCACCT 555
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

795 GGAACGGCAGATGGGCTACTATCTGATTGATGATGATGATGATGATGATGATGATGATGAT 854
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```



```

Db      556 GGAGGACAGATGGGCTACTGATCCAGATGTATACATCCCTCGCTCATCGTCAT 615
QY      855 CTTGCTCTGGGCTCTCTTCTGGATCAACATGATGCTGCGCTCGCC 901
Db      616 CTTGCTCTGGGCTCTCTTCTGGATCAACATGATGAGCGCGCGCGCC 662

RESULT 14
LOCUS   CD353729
DEFINITION UI-M-GMO-csc-m-22-0-UI.r1 NIH_BMAP_GMO Mus musculus cDNA clone
IMAGE:30360981 5', mRNA sequence.
CD353729
CD353729.1 GI:31146230
EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 805)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousef1.html
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

Seq primer: pyX-5.
Location/Qualifiers
1. 805
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30360981"
/tissue_type="whole brain"
/dev_stage="1, 5 and 15 days newborn"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_GMO"
/notes="Organ: Brain; Vector: pyX-Asc; Site:1: Ecor I;
Site:2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was size
selected according to mRNA size fraction, ligated with Ecor
I adaptor, digested with NotI and then cloned
directionally into pyX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is CGAACGCAAT. This library was created for the University
of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

ORIGIN
Query Match 24.2%; Score 397; DB 6; Length 805;
Best Local Similarity 71.9%; Pred.No.2.3e-99;
Matches 536; Conservative 0; Mismatches 200; Indels 9; Gaps 1;

QY      663 TGTCTCTGCTGCCAAGTGGCTGAGGGGCTGACTCTGCCCGGATTTATCTGCGGATGA 722
Db      3 TGATGCTCCAGTACAGATGTTGCTGAAGGACTCACCTTGCCCGGATTTATTTGAAGAAGA 62
QY      723 GAAGGATCTAGGCTGTTGTACCAAGCACTACACAGGGAATTCACCTGCATCGAGGT 782
Db      63 GAAGGAGCTTGGTTATTGCAAGCAATTAACACACTGGCAAGTTTACCTGCATTGAGGT 122

```

```

QY      783 AAAGTTTCACTGGAACGCGCAGATGGGCTACTATCTGATTCAGATGTATATCCCGAGCCT 842
Db      123 CAAAGTTTCACTGGAGCGCCAGATGGGCTACTATTTGATCCAGATGTATATCCCGAGCCT 182
QY      843 ACTCATCGTCAATCCTGCTGGGCTCTCTTCTGGATCAACATGGATGCTGCGCCGCG 902
Db      183 GTTGATAGTCATTTTGTCTCTGGGCTCTCTTTTGGATAAAACATGGATGCGAGCCCTGCCAG 242
QY      903 TGTGGGCTGGGCATCACACCGTCTCACCATGACCAACCCAGAGCTCTGGCTCCCGGCG 962
Db      243 GGTTCGCCCTGGGCATCAACACAGTCTCTGCAATGACTACACAGATTCAGGTTTCAGGGCG 302
QY      963 CTCTTTGCTTAAGGTGTCTTACGTGAAGGCAATGCACATCTCGATGGCTGTGTCTGCT 1022
Db      303 ATCTCTGCCAAAGGTCTCTTATGTGAAGCAATTCACATCTGATGGCGGTATGCTTCT 362
QY      1023 CTTTGTGTCGCTGCTGCTGCTGAGTATGCTGCCATAAATTTTCTTCGTCAGCATAA 1082
Db      363 CTTTGTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 422
QY      1083 AGAATTCATACACTTCGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1142
Db      423 GGAGTTCCTTCTGCTCTCCGAGACGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 482
QY      1143 AGAAGTTCCTTCTTATTTCCGTGGCTATGGCTTGGGCCACTGCTCGAGGCAAGAGATGG 1202
Db      483 TGAAGTTCCTTCTTAACTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 542
QY      1203 AGGTCCAAATGGAAGTTCGTCATTTATAGTCCCAACCTCCAGCCCTCTCTTCTAAGGA 1262
Db      543 CACAGCTGTCAAGGCTAC-----ACCTGCCAACCCACTTCCACAAACCCCAAGA 593
QY      1263 AGGGAACCCGCGGAACTCTAGTGGACTGACCGAGAGAAATGACACCATCTCCCG 1322
Db      594 TGCAGATGCTATCAAGAAGAGTTTGTGGATCGGGCAAAAGAAATGACACCATCTCTCG 653
QY      1323 GGCTGTCTTCCCTTTCACCTTCTCTCATCTTCAATATCTTCTACTGGGTTGTCTATAAAGT 1382
Db      654 AGCTGCTTCCCACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 713
QY      1383 GCTATGGTCAGAAGATATCCACCAG 1407
Db      714 CATTGGCATGAAGATGTCACAAG 738

RESULT 15
LOCUS   CN266978
DEFINITION CN266978
ACCESSION CN266978
VERSION   CN266978.1 GI:47283392
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 540)
Branderberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Flak,G.J.,
Li,Y., Xu,C., Fang,R., Guegl,K., Rao,M.S., Mandalam,R.,
Lebkowski,J and Stanton,L.W.
Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)
Contact: Branderberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbranderberger@geron.com
Insert Length: 540 Std Error: 0.00.
Location/Qualifiers

FEATURES

```



```

source      1. .540
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /tissue_type="embryonic stem cells, embryoid bodies
            derived from H1, H7 and H9 cells"
            /clone_lib="GRN EB"
            /note="oligo dT primed, full-length enriched cDNA library
            from embryoid body outgrowths derived from hES cell lines
            H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
            conditions."

ORIGIN
Query Match      23.1%; Score 378.6; DB 7; Length 540;
Best Local Similarity 99.0%; Pred. NO. 2.9e-94;
Matches 381; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 66 CCTCAGGGTGGCTTGGCAAAAGAGGAAGTCAAAATCTGGAACCAAGGGGTCCAGGCCAT 125
Db 156 CCCAGGGGTGGCTTGGCAAAAGAGGAAGTCAAAATCTGGAACCAAGGGGGCCAGGCCAT 215

Qy 126 GTCCCTCTGATTTCTAGACAAACTTATGGGGCGAACATCTGGATATGATGCCAGGAT 185
Db 216 GTCCCTCTGATTTCTAGACAAACTTATGGGGCGAACATCTGGATATGATGCCAGGAT 275

Qy 186 TCGGGCCCAATTTAAAGGGCCACCCGTGAACGTGACCTGCAACATCTTCATCAACAGTTT 245
Db 276 TCGGGCCCAATTTAAAGGGCCACCCGTGAACGTGACCTGCAACATCTTCATCAACAGTTT 335

Qy 246 CAGCTCCGTCAACCAAGACCAATGACTACCGGGTGAATGCTCTTGGCGCAACAGTG 305
Db 336 CAGCTCCGTCAACCAAGACCAATGACTACCGGGTGAATGCTCTTGGCGCAACAGTG 395

Qy 306 GAATGACCCACGCTCTCTTACCGAGAATATCTGATGACTCTCTGGACCTCGATCCCTC 365
Db 396 GAATGACCCACGCTCTCTTACCGAGAATATCTGATGACTCTCTGGACCTCGATCCCTC 455

Qy 366 CATGCTGGACTCTATCTGGAAGCCAGACCTCTTCTTTGCTATGAGAAAGGGGCCAATT 425
Db 456 CATGCTGGACTCTATCTGGAAGCCAGACCTCTTCTTTGCTATGAGAAAGGGGCCAATT 515

Qy 426 CCATGAGGTGACCAAGCAACCAAG 450
Db 516 CCATGAGGTGACCAAGCAACCAAG 540

```

Search completed: April 28, 2005, 16:13:05  
Job time : 6708.22 secs



Result No.	Query %			DB	ID	Description
	Score	Match	Length			
1	44	91.7	39796	6	AX037570	Sequence
C 2	44	91.7	39796	6	HSU3563	Z33848 Human DNA s
3	28.4	59.2	194225	2	AC137157	AC137157 Mus muscu
4	28.4	59.2	200858	10	AC118255	AC118255 Mus muscu
5	26.6	55.4	110000	2	AC097996.2	Continuation (3 of
C 6	26.6	55.4	175846	2	CR759777	CR759777 Danio rer
C 7	26.6	55.4	222983	2	AC129695	AC129695 Rattus no
C 8	26.6	55.4	233746	2	AC099111	AC099111 Rattus no
C 9	26.6	55.4	253973	2	AC094544	AC094544 Rattus no
C 10	26.6	55.4	259833	2	AC123469	AC123469 Rattus no
11	26.6	55.4	283158	2	AC137193	AC137193 Rattus no
C 12	26.4	55.0	129968	9	HS24985	AL035691 Human DNA
C 13	26.4	55.0	137436	9	AF348209	AF348209 Homo sapi
C 14	26.4	55.0	138376	9	AF069333	AF069333 Homo sapi
C 15	26.4	55.0	139947	9	AY293855	AY293855 Homo sapi
C 16	26.2	54.6	594	11	BV155969	BV155969 PRAMSECO
17	26.2	54.6	596	11	BV098593	BV098593 PRAMSECO
C 18	26.2	54.6	94535	10	RN141K6	AL603724 Rattus no
C 19	26.2	54.6	102358	10	AL645479	AL645479 Mouse DNA

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 39796)  
 BURGESS, J.  
 Direct Submission  
 Submitted (05-MAR-2003) Wellcome Trust Sanger Institute, Hinxton,  
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
 humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
 ----- Genome Center  
 Center: Wellcome Trust Sanger Institute  
 Center code: SC  
 Web site: <http://www.sanger.ac.uk>  
 Contact: [humquerry@sanger.ac.uk](mailto:humquerry@sanger.ac.uk)  
 -----  
 During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence submission  
 corresponding to the overlapping clone, as we submit sequences with  
 only a small overlap as described above.  
 This sequence was finished as follows unless otherwise noted: all  
 regions were either double-stranded or sequenced with an alternate  
 chemistry or covered by high quality data (i.e., phred quality >= 30);  
 an attempt was made to resolve all sequencing problems, such as  
 as compressions and repeats; all regions were covered by at least  
 one plasmid subclone or more than one M13 subclone; and the  
 assembly was confirmed by restriction digest, except on the rare  
 occasion of the clone being a YAC.  
 The following abbreviations are used to associate primary accession  
 numbers given in the feature table with their source databases:  
 Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information  
 on the WORMPEP database can be found at  
[http://www.sanger.ac.uk/projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/projects/C_elegans/wormpep) This sequence  
 was generated from part of bacterial clone contigs of human  
 chromosome X, constructed by the Sanger Centre Chromosome X Mapping  
 Group. Further information can be found at  
<http://www.sanger.ac.uk/HGP/ChrX>  
 LLOXNC01-35G3 is from the Lawrence Livermore National Laboratory  
 flow-sorted  
 X chromosome cosmid library LLOXNC01  
 VECTOR: lawr16,  
 Location/Qualifiers  
 1..39796  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="RZPD:LLNC110G0335"  
 /db\_xref="taxon:9606"  
 /chromosome="X"  
 /clone="LLOXNC01-35G3"  
 /clone\_lib="LLOXNC01"  
 -----  
 ORIGIN  
 Query Match 91.7%; Score 44; DB 9; Length 39796;  
 Best Local Similarity 100.0%; Pred. No. 4e-06;  
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 5 CATCCATCTCTGCAGCCCTCTGCCATCTCTGTCACCTTCAGTT 48  
 ||||||||||||||||||||||||||||||||||||||||||||  
 Db 4077 CATCCATCTCTGCAGCCCTCTGCCATCTCTGTCACCTTCAGTT 4034  
 -----  
 RESULT 3  
 AC137157  
 LOCUS  
 DEFINITION Mus musculus chromosome 8 clone RP23-377M11, WORKING DRAFT  
 SEQUENCE, 9 unordered pieces.  
 AC137157  
 AC137157.2 GI:52077837  
 HTG: HTGS\_PHASE1; HTGS DRAFT; HTGS\_ACTIVEFIN.  
 KEYWORDS Mus musculus (house mouse)  
 SOURCE  
 ORGANISM  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 194225)  
 WILSON, R.K.  
 TITLE The sequence of Mus musculus clone  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 194225)  
 McPherson, J.D. and Waterston, R.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (18-NOV-2002) Genome Sequencing Center, 4444 Forest Park  
 Parkway, St. Louis, MO 63108, USA  
 REFERENCE 3 (bases 1 to 194225)  
 WILSON, R.K.  
 TITLE Direct Submission  
 JOURNAL Submitted (15-SEP-2004) Genome Sequencing Center, 4444 Forest Park  
 Parkway, St. Louis, MO 63108, USA  
 COMMENT On Sep 15, 2004 this sequence version replaced gi:25054283.  
 -----  
 Center: Genome Center  
 Center: Washington University Genome Sequencing Center  
 Center code: WUGSC  
 Web site: <http://genome.wustl.edu>  
 Contact: [submissions@watson.wustl.edu](mailto:submissions@watson.wustl.edu)  
 ----- Project Information -----  
 Center project name: M\_BA0377M11  
 ----- Summary Statistics -----  
 Sequencing vector: M13; 0%  
 Chemistry: Dye-primer ET; 0% of reads  
 Assembly: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990319  
 Consensus quality: 190423 bases at least Q40  
 Consensus quality: 190874 bases at least Q30  
 Consensus quality: 191269 bases at least Q20  
 -----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 9 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \* 1 1386: contig of 1386 bp in length  
 \* 1387 1486: gap of unknown length  
 \* 1487 2690: contig of 1204 bp in length  
 \* 2691 2790: gap of unknown length  
 \* 2791 6363: contig of 3573 bp in length  
 \* 6364 6463: gap of unknown length  
 \* 6464 18259: contig of 11796 bp in length  
 \* 18260 18359: gap of unknown length  
 \* 18360 35861: contig of 17502 bp in length  
 \* 35862 35962: gap of unknown length  
 \* 35962 59175: contig of 23214 bp in length  
 \* 59176 59275: gap of unknown length  
 \* 59276 81565: contig of 22290 bp in length  
 \* 81566 81666: gap of unknown length  
 \* 81666 127719: contig of 46054 bp in length  
 \* 127720 127819: gap of unknown length  
 \* 127820 194225: contig of 66406 bp in length.  
 -----  
 FEATURES  
 source  
 1..194225  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /chromosome="8"  
 /clone="RP23-377M11"  
 1..1386  
 /note="assembly\_name:Contig14"  
 1487..2690  
 /note="assembly\_name:Contig25"  
 2791..6363  
 /note="assembly\_name:Contig26"  
 6464..18259  
 /note="assembly\_name:Contig27"

```

/notes="assembly_name:Contig27"
18360..35861
/notes="assembly_name:Contig28
clone_end:17
vector_side:right"
35962..59175
/notes="assembly_name:Contig29"
59276..81565
/notes="assembly_name:Contig30"
81666..127719
/notes="assembly_name:Contig31"
127820..194225
/notes="assembly_name:Contig32
clone_end:SP6
vector_side:right"

ORIGIN
Query Match      59.2%; Score 28.4; DB 2; Length 194225;
Best Local Similarity 76.1%; Pred. No. 3.1;
Matches 35; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 3 CTATCATCTCTGAGCGCTCTGCACCTCTGTCACCTCTGTCAGTT 48
|||||
Db 7117 CTCTCCAGACTCAGCAGGCTCTGCACCTCTCTCTCTCTCTCAGTT 7162

RESULT 4
AC118255 200858 bp DNA linear ROD 09-SEP-2004
DEFINITION Mus musculus chromosome 8, clone RP24-89A4, complete sequence.
ACCESSION AC118255
VERSION AC118255.32 GI:51948654
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Birren,B., Nussbaum,C. and Lander,E.
1 (bases 1 to 200858)
Mus musculus chromosome 8, clone RP24-89A4
Unpublished
2 (bases 1 to 200858)
Birren,B., Linton,L., Nussbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Collangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Landers,T., Lehocsky,J., Levine,R., Lindblad-Toh,K., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Riback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (14-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 200858)
Birren,B., Nussbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
Bloom,T., Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J.,
Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,
DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D.,
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
Landers,T., Lehocsky,J., Levine,R., Lindblad-Toh,K., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (16-JUL-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 200858)
Birren,B., Nussbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
Bloom,T., Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J.,
Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,
DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D.,
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (09-SEP-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 9, 2004 this sequence version replaced gi:50345206.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence submissions@broad.mit.edu
----- Project Information
Center project name: L24331
Center clone name: 89_A_4
-----
Location/Qualifiers
1..200858
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosomes="8"
/maps="8"
/clone="RP24-89A4"
/clone_lib="RP24-89A4"
1..4
/note="clone boundary
clone end:SP6
site:MboI"
684..937
/rpt_family="B4"
938..968
misc_feature
repeat_region
repeat_region

```



Consensus quality: 173690 bases at least Q40  
 Consensus quality: 174041 bases at least Q30  
 Consensus quality: 174473 bases at least Q20  
 Insert size: 175146; sum-of-contigs  
 Quality coverage: 8.97x in Q20 bases; sum-of-contigs Quality coverage: 8.44x in Q20 bases; agarose-fp

-----  
 \* NOTE: This is a 'working draft' sequence. It currently consists of 8 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 2342: contig of 2342 bp in length  
 \* 2342: gap of 100 bp  
 \* 2443 25852: contig of 23410 bp in length  
 \* 25853 25922: gap of 100 bp  
 \* 25953 97504: contig of 71552 bp in length  
 \* 97505 97604: gap of 100 bp  
 \* 97605 118987: contig of 21383 bp in length  
 \* 118988 119087: gap of 100 bp  
 \* 119088 150462: contig of 31375 bp in length  
 \* 150463 150562: gap of 100 bp  
 \* 150563 156801: contig of 6239 bp in length  
 \* 156802 156901: gap of 100 bp  
 \* 156902 163051: contig of 6150 bp in length  
 \* 163052 163151: gap of 100 bp  
 \* 163152 175846: contig of 12695 bp in length.

## FEATURES

## source

1. 175846  
 /organism="Danio rerio"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:7955"  
 /clone="DKFZ-162P5"  
 /clone\_lib="Daniokey"

## misc\_feature

1. 2342  
 /note="assembly\_fragment:00027"

## misc\_feature

2443..25852  
 /note="assembly\_fragment:01017"

## misc\_feature

25953..37504  
 /note="assembly\_fragment:01383"

## misc\_feature

97605..118987  
 /note="assembly\_fragment:00369"

## misc\_feature

119088..150462  
 /note="assembly\_fragment:00670"

## misc\_feature

150563..156801  
 /note="assembly\_fragment:00059"

## misc\_feature

156902..163051  
 /note="assembly\_fragment:00104"

## misc\_feature

163152..175846  
 /note="assembly\_fragment:00182.0"

## ORIGIN

Query Match 55.4%; Score 26.6; DB 2; Length 175846;  
 Best Local Similarity 78.0%; Pred. No. 16;  
 Matches 32; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 5 CATCCATCTCTGACGCTCTGCCATCTCTGCACTTCA 45  
 |||||  
 DB 170401 CATTCATAATCAGAGCTCTGCCATTTGTATCACTCTAA 170361

## RESULT 7

AC129695/c

## LOCUS

222983 bp DNA linear HTG 10-MAY-2003

## DEFINITION

Rattus norvegicus clone CH230-11B16, WORKING DRAFT SEQUENCE, 2  
 unordered pieces.

## ACCESSION

AC129695

## VERSION

HTG; HTGS PHASE1; HTGS DRAFT; HTGS\_FULLTOP.

## KEYWORDS

Rattus norvegicus (Norway rat)

## SOURCE

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

## REFERENCE

1 (bases 1 to 222983)

## AUTHORS

Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Ayagi, A., Ayodeji, M., Baca, S., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falle, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Geisregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hoques, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshuwa, L., Loulseg, H., Lozano, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangun, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mathew, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwokedemeh, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Prannkoch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Saverly, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajls, D., Speed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willison, R., Wlezyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

## TITLE

Direct Submission

## JOURNAL

Unpublished

## REFERENCE

2 (bases 1 to 222983)

## AUTHORS

Worley, K. C.

## TITLE

Direct Submission

## JOURNAL

Submitted (01-AUG-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

## REFERENCE

3 (bases 1 to 222983)

## AUTHORS

Rat Genome Sequencing Consortium.

## TITLE

Direct Submission

## JOURNAL

Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One





TITLE JOURNAL	Direct Submission				RESULT 9			
	Submitted (09-NOV-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA				AC094544/c			
	3 (bases 1 to 233746)				LOCUS			
	Rat Genome Sequencing Consortium.				DEFINITION			
REFERENCE AUTHORS TITLE JOURNAL	Direct Submission				AC094544			
	Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA				AC094544, 7			
	On May 10, 2003 this sequence version replaced gi:24819632. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.				AC094544, 7			
	Genome Center Center: Baylor College of Medicine Center code: BCM Web site: http://www.hgsc.bcm.tmc.edu/ Contact: hgsc-help@bcm.tmc.edu Project Information Center project name: GGPT Center clone name: CH230-171D15 Summary Statistics Assembly program: Atlas 3.0; Consensus quality: 223529 bases at least Q40 Consensus quality: 225655 bases at least Q30 Consensus quality: 227116 bases at least Q20 Estimated insert size: 232984; sum-of-contigs estimation Quality coverage: 8x in Q20 bases; sum-of-contigs estimation				AC094544, 7			
COMMENT	NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).				AC094544, 7			
	NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.				AC094544, 7			
	1 228528: contig of 228528 bp in length				AC094544, 7			
	228529 228628: gap of unknown length				AC094544, 7			
FEATURES	228629 229841: contig of 1213 bp in length				AC094544, 7			
	229842 229941: gap of unknown length				AC094544, 7			
	229942 233746: contig of 3805 bp in length.				AC094544, 7			
	Location/Qualifiers				AC094544, 7			
misc_feature	1. 233746				AC094544, 7			
	/organism="Rattus norvegicus"				AC094544, 7			
	/mol_type="genomic DNA"				AC094544, 7			
	/db_xref="taxon:10116"				AC094544, 7			
ORIGIN	/clone="CH230-171D15"				AC094544, 7			
	1. 1039				AC094544, 7			
	/note="wgs_contig"				AC094544, 7			
	Query Match 55.4%; Score 26.6; DB 2; Length 233746;				AC094544, 7			
QY	Best Local Similarity 78.0%; Pred. No. 17;				AC094544, 7			
	Matches 32; Conservative 0; Mismatches 9; Indels 0; Gaps 0;				AC094544, 7			
	4 TCATCCATCTCTGCGCCCTTCCCATCTCTCTGCTTC 44				AC094544, 7			
					AC094544, 7			
Db	202426 TCATCCATCTCTTCAACCCCTCCCATCTATCATCTC 202386				AC094544, 7			
					AC094544, 7			
					AC094544, 7			
					AC094544, 7			

RESULT 9  
 AC094544/c  
 LOCUS  
 DEFINITION  
 AC094544  
 AC094544, 7  
 GI:30466710  
 HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_ENRICHED.  
 Rattus norvegicus (Norway rat)  
 Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus  
 1 (bases 1 to 233746)  
 Muzny, D.Marie, Metzker, M.Lee., Abramson, S., Adams, C., Alder, J.,  
 Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,  
 Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,  
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,  
 Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,  
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,  
 Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,  
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,  
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,  
 Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,  
 Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Diya, K.,  
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,  
 Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,  
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,  
 Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,  
 Gregorogis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,  
 Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,  
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,  
 Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M.,  
 Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,  
 Karpinsky, S., Kelly, S., Khan, Z., King, L., Kovar, C.,  
 Kowalski, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,  
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,  
 Lorensheuer, L., Loulseghe, H., Lozano, R.J., Lu, X., Ma, J.,  
 Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,  
 Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E.,  
 Mawney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,  
 Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,  
 Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Naif, L.,  
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,  
 Nwankwelu, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,  
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,  
 Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.,  
 Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,  
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,  
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,  
 Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,  
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajda, D.,  
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Soosa, J.,  
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,  
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,  
 Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,  
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,  
 Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K.,  
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, V.,  
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von  
 Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,  
 Weinstock, G. and Gibbs, R.A.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 233746)  
 Worley, K.C.  
 Direct Submission  
 Submitted (15-SEP-2001) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 233746)  
 Rat Genome Sequencing Consortium.

**TITLE**  
JOURNAL

**COMMENT**

Direct Submission  
Submitted (09-MAY-2003) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On May 9, 2003 this sequence version replaced gi:24942135.  
The sequence in this assembly is a combination of BAC based reads  
and whole genome shotgun sequencing reads assembled using Atlas  
(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described  
in the feature table below represents a scaffold in the Atlas  
assembly (a 'contig-scaffold'). Within each contig-scaffold,  
individual sequence contigs are ordered and oriented, and separated  
by sized gaps filled with Ns to the estimated size. The sequence  
may extend beyond the ends of the clone and there may be sequence  
contigs within a contig-scaffold that consist entirely of whole  
genome shotgun sequence reads. Both end sequences and whole genome  
shotgun sequence only contigs will be indicated in the feature  
table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
----- Project Information  
Center project name: CH230-4118  
Center clone name: CH230-4118  
----- Summary Statistics  
Assembly program: Atlas  
Consensus quality: 229658 bases at least Q40  
Consensus quality: 232789 bases at least Q30  
Consensus quality: 235199 bases at least Q20  
Estimated insert size: 238134; sum-of-contigs estimation  
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 3 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\*  
\* 1 251012: contig of 251012 bp in length  
\* 251013 251112: gap of unknown length  
\* 251113 252217: contig of 1105 bp in length  
\* 252218 252317: gap of unknown length  
\* 252318 253973: contig of 1656 bp in length.  
\* Location/Qualifiers  
\* 1..253973  
\* /organism="Rattus norvegicus"  
\* /mol\_type="genomic DNA"  
\* /db\_xref="taxon:10116"  
\* /clone="CH230-4118"  
\* 1..1024  
\* /note="wgs contig"  
\* 5428..6613  
\* /note="wgs contig"  
\* 36648..37724  
\* /note="wgs contig"  
\* 95090..96935  
\* /note="wgs contig"  
\* 245711..246611  
\* /note="clone\_boundary"  
\* clone\_end:17  
\* site:EcoRI  
\* end\_sequence:BH307674"  
\* 249761..251012  
\* /note="wgs end\_extension"  
\* clone\_end:17"

**FEATURES**  
source  
misc\_feature  
misc\_feature  
misc\_feature  
misc\_feature  
misc\_feature  
misc\_feature  
misc\_feature  
Query Match

**ORIGIN**  
JOURNAL

**Best Local Similarity** 78.0%; Pred. No. 17;  
Matches 32; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 2 GCTCATCCATATCTGCAGCCCTCTGCCATCTCTGTCACCTT 42  
Db 99678 GCTAGTCCAGAAATTACAGATATCTGCCATCTCTGTCACCTT 99638

**RESULT 10**  
AC123469 259833 bp DNA linear HTG 19-NOV-2002  
Rattus norvegicus clone CH230-5003.6, WORKING DRAFT SEQUENCE, 2  
unordered pieces.  
AC123469  
VERSION  
AC123469.4 GI:25086546  
HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.  
KEYWORDS  
Rattus norvegicus  
SOURCE  
Rattus norvegicus (Norway rat)  
ORGANISM  
Mammalia; Eutheria; Rodentia; Chordata; Vertebrata; Euteleostomi;  
Eukaryota; Metazoa; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus

**REFERENCE**  
1 (bases 1 to 259833)  
**AUTHORS**  
Muzny, D. Marie, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J.,  
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,  
Anyalebechi, V., Ayagi, A., Ayodeji, M., Baca, E., Baden, H.,  
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,  
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,  
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,  
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,  
Chacko, J., Chavez, D., Chen, R., Chen, Y., Chen, Z., Chu, J.,  
Cleveland, C., Cockrell, R., Cox, C., Coyly, M., Cree, A., D'Souza, L.,  
Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,  
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,  
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,  
Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,  
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,  
Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,  
Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,  
Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,  
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,  
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hoggins, M.,  
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,  
Jacks, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,  
Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,  
Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,  
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,  
Lorensunewa, L., Louie, H., Lozano, R. J., Lu, X., Ma, J.,  
Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,  
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,  
Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,  
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,  
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Naik, L.,  
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,  
Nwankwelu, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,  
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,  
Plopper, F., Poinexter, A., Popovic, D., Primus, E., Pu, L.,  
Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reich, R.,  
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, P.,  
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,  
Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,  
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajs, D.,  
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,  
Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,  
Taylor, T., Thomas, R., Thomas, S., Tingey, A., Trejos, Z., Umani, K.,  
Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,  
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,  
Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K.,  
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,  
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., Von  
Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,  
Weinstock, G. and Gibbs, R. A.

**TITLE**  
JOURNAL

**Query Match** 55.4%; Score 26.6; DB 2; Length 253973;

```

REFERENCE
AUTHORS
TITLE
JOURNAL

2 (bases 1 to 259833)
Worley, K.C.
Direct Submission
Submitted (29-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 259833)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (19-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

On Nov 19, 2002 this sequence version replaced gi:23265401.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GXP2016
Center clone name: CH230-50016
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 241805 bases at least Q40
Consensus quality: 245039 bases at least Q30
Consensus quality: 247120 bases at least Q20
Estimated insert size: 247152; sum-of-ctnigs estimation
Quality coverage: 7x in Q20 bases; sum-of-ctnigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.
* 1 258704: contig of 258704 bp in length
* 258705 258804: gap of unknown length
* 258805 259833: contig of 1029 bp in length.

FEATURES
source
Location/Qualifiers
1. .259833
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-50016"

misc_feature
1. .1303
/notes="wgs end_extension
clone_end:Sp6"
6179. .6809
/notes="clone boundary
clone_end:Sp6
site:

misc_feature
end sequence:BH336040"
complement(24927. .257242)
/notes="clone boundary
clone_end:T7
site:

misc_feature
end sequence:BH336039"

```

```

misc_feature
68765. .70448
/notes="wgs end_extension
clone_end:T7"

ORIGIN

Query Match 55.4%; Score 26.6; DB 2; Length 259833;
Best Local Similarity 78.0%; Pred. No. 17;
Matches 32; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 4 TCATCATCTCTGCGAGCCCTCTGCGATCTCTGTCACCTTC 44
|||||
DB 157894 TCATCATCTCTCTCAACCCCTCCACATCTATCCTTC 157854
|||||

RESULT 11
AC137193 283158 bp DNA linear HTG 19-NOV-2002
Rattus norvegicus clone CH230-unknown, *** SEQUENCING IN PROGRESS
*** 4 unordered pieces.
AC137193
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
Rattus norvegicus (Norway rat)
Rattus norvegicus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

1 (bases 1 to 283158)
Muzny, D. Marie, Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Ayagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A.,
Chacko, J., Chavez, D., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, B., Kovar, C.,
Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorensu, L., Loulseg, H., Lozano, R. J., Lu, X., Ma, J.,
Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhinney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwaokeme, O., Okwuonu, G., Olarnpungoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plopper, F., Polindexter, A., Popovic, D., Primus, E., Pu, L.,
Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,
Sanders, W., Savary, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajls, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Stearle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wlarczyk, R., Wooden, H., Worley, K.,

```



```
Db 592 AGGTGCTGATATCTGTTGCCTTCTGCATGACTGTAAGTTTC 635
|||||
RESULT 13
AF348209/c 137436 bp DNA linear PRI 27-MAY-2001
LOCUS Homo sapiens IGF2R gene, complete cds.
DEFINITION AF348209.1 GI:14210065
ACCESSION AF348209.1
VERSION AF348209.1
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 137436)
Human M6P/IGF2R complete genomic sequence
Unpublished
2 (bases 1 to 137436)
Killian, J.K.
Direct Submission
Submitted (10-FEB-2001) Pathology, Duke University, USA
Location/Qualifiers
1. 137436
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/map="6q26-q27"
join(2..296,22079..22218,39906..40030,41583..41681,
53471..55603,58084..58213,60449..60554,63450..63612,
63841..64006,65318..65421,71444..71608,74032..74172,
73398..75541,76629..76766,77382..77529,78043..78220,
78676..78791,79259..79427,81357..81536,87308..87409,
88907..89008,89790..89982,91431..91601,92387..92530,
92637..92812,93416..93503,94299..94514,95285..95415,
95688..95785,99133..99269,100752..100942,102794..102920,
103649..103768,104097..104353,104641..104859,
106731..106880,109085..109246,110465..110672,
111014..111160,114835..115069,115880..116016,
118918..119032,119992..120138,120801..120988,
127324..127510,133404..133556,134631..134700,
135559..137436)
/product="IGF2R"
join(148..296,22079..22218,39906..40030,41583..41681,
53471..55603,58084..58213,60449..60554,63450..63612,
63841..64006,65318..65421,71444..71608,74032..74172,
73398..75541,76629..76766,77382..77529,78043..78220,
78676..78791,79259..79427,81357..81536,87308..87409,
88907..89008,89790..89982,91431..91601,92387..92530,
92637..92812,93416..93503,94299..94514,95285..95415,
95688..95785,99133..99269,100752..100942,102794..102920,
103649..103768,104097..104353,104641..104859,
106731..106880,109085..109246,110465..110672,
111014..111160,114835..115069,115880..116016,
118918..119032,119992..120138,120801..120988,
127324..127510,133404..133556,134631..134700,
135559..137436)
/feature="IGF2R"
/codon_start=1
/product="IGF2R"
/protein_id="AAK56918.1"
/db_xref="GI:14210066"
/translation="MGAAGRSPHLGPARRPQRSLQLQLLLVAAPGTSQAQAP
FPELCSYWEADVTKNVLKINICGSDIVQCGSSAVCMHDKTRTHSVGDSVLR
SATRSLEFNTVSCDQGNHRVQSIAFLCKLTGTEFVTECEVHFWRITAA
CKKIDIPKANKVPCYVDFEELRKHDLNPLIKSGAYLVDDSDPDSLSFINVCRDITL
RFGSQLRACRACLVGRHQAFDVGQFDRGLKVRDLRLVLYVREAEAGLDFCD
GHSPTVITFVCPSEERRETIPLKATSKNCRVIEIWTYACHRDYLSKTCSLGSEQ
QVSDILTPLAQSGLSSYSDGKYLFLYNVCGETIOPCNKKQAAVCVKKSDTSOV
KAAGRHNQTLRYSDGDLILIFPGDECSGFGFMSVFNPCNKTAGNKGKTPFTG
EVDCTFTFTYAIYACVKEKEDLLCGATDGKRYDLSALVRHAEPEQNWEAVDSQTE
```

TEKHHFFINICHRVLOEGKARGCPEDAAVCAVDKNGSKNLGKFISSPMKEKGNOLSY  
SDGDCGCHKKIKTNITLVCKFGDLESAVPLKTSRGEGCFYEFWRHTAAACVLKSTEG  
ENCTVDFDSQAGFSFSLSPKNGAYKTKYDFYINVCQPVSPVSPQPDSDGACQVA  
KDEKTNLGNLNAKLSYDGMIOQLNRYGGTTPYNNERTPRATLITFLCDRAGVGF  
EYDEEDNSTYFNRYTSTACPEPELVCTDPTSTEQYDLSSLAKEGGLGNGWYAMD  
NVEGEWTRKYIYINVCPLNVPVGCNRYASACOMKYKQDQGFTEVUSTSLNLMKATG  
PVEDSGSLLEYVANGSACTTSDGRTQVTTTRILHVCSTRGLNSHP1FSLNVECVUSP  
LANTEAACPIQTITTDQACSTRDPNSGVFNLPNLSQGVNYSIGIKFPHNVCGT  
MPVCGTILGKAPGCAEATQTBELKNKPARVPIEIKSLQSTEGTITTYGKPLSAK  
GTADAFIVFVNCDDYVSGFLFHQIDISGQIRNTYEFETALACVSPVDCQVTD  
LAGNEDYLTGLSTVRKPTAVDTISVDYGRKTFYLSVNCNPLPIPGCQGSAPVQLODG  
GNSWNLGVVMSPOQAANGSLSIMYVNGDKCQNRSTRTPECAOISGSAPVQLODG  
CEYVTRTVEACPVVRVGDNCEVKDPBHNLYDLKPLGLNDTIVSAGEYTVYFVVC  
GXLSDDVCFSTDKSVKVSQCKRPEQGFHVLVAGLLTQKLYENGLLLKNFTGCDTCH  
KYQRTSLFFYCDRGTQRPFLKTSYFELWRTYACPPFDTECSFADGKGNV  
FDLSLSRYSDNWEIITGVDPHEYLINVCYKLAQAGTPECPPEAAACLLGSKGVN  
LGRVRDGPQWRDGIIVLYKVDGLDPCDGIKTKSTTIRFTCSQVNSRPMFISAVDC  
EYTFWAPTATACPMKSNHDDCQVNTNPSHGLFDLSLSGRAGTAAAYSEKGLVYMSI  
CGENENCPPGVGACFCGQTRPISYGKANKRLRYVDVQLVYKDGSPCPSKSLYSKSVI  
SFVCPPEARPTNRPMLISLDKQCTILFFSWHTPLACEQATECSVRNGSSIVDUSPLIH  
RTGYEADVEDSDDADTNPDFYINICQPLNPMHGVPCPAGAAVCKVPIGDPIDIGR  
VAGPPIPLNIAEYILNFESFPCADKHNTSLIAFHCKRGVSGTKLRTSECD  
FVPEWETPVVCPDEVMDGCTLTDEQLYSFNLSSLSTSTFKVTRDSRYVSGVCTFA  
VGPEOGCKDGGVCLLSGKTGASFGRLQSMKLDYRHQDRAVLVYVNGDRCPPTDDG  
VPCVFPPIFGKSYSECIIESRAKLCWSTADYDRDHENGFCRHSNYSYTSIIFKCD  
EDEDIGRPQVSEVRGCDVTFEWKTKVCPCKLECKYVQKHKTLYDLRLSLSSUTGSW  
LVHNGVSYINLCQIKYKPLGFLSERASICRRTTGGVQLVGLVHTQKLVIGDKVWV  
TSGYKPGCGNKTSASVBLDTKTVPRAFPRFDIDSTYTFSDSRAACAVKPOEV  
QWNGTITNPNGKISGLDIYKPLFRASGDMRTNGDNYLIEIQLSSIYSSRNAPCSG  
ANICQVKNDOHFSRKYSDKTYLOQDGLDVVPASSSSCKGKDKTKSVSTIIFPHC  
DPLVEDGPIPEFSHETADCOYLFSTWTSVAVCPGLGVGDFSDNPGDGMKHLGSRQAV  
GAVLSLLVALUCCLLALLYKKERRETVISKLTTCRNSNNVSYGKVKRKEETDE  
NETEWLMESIQLPFRPQKGEQGHITTKSVKALSSHLGDDQDSEDLVITPEVXVF  
SGRGAADSSHEFVRNAQSNALQEREDDRVGLVRGEKARKGKSSSAQKTVSTKLVSF  
HDDSDDELLHI"

ORIGIN

Query Match 55.0%; Score 26.4; DB 9; Length 137436;  
Best Local Similarity 75.0%; Pred. No. 19;  
Matches 33; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 AGTCATCCATCTCTGCGACCTCTGCGCATCTCTGCTCACTTTC 44  
Db 43391 AGGTGCTGATATCTGTTGCCTTCTGCATGACTGTAAGTTTC 43348

AF069333 138376 bp DNA linear PRI 17-SEP-2001  
Homo sapiens insulin-like growth factor II receptor (IGF2R) gene,  
partial cds.  
AF069333 AF007273 AF069334 AF069335 AF069336 AF069337 AF069338  
AF069339 AF069340 AF069341 AF069342 AF069343 AF069344 AF069345  
AF069346 AF069347 AF069348 AF069349 AF069350 AF069351 AF069352  
AF069353 AF069354 AF069355 AF069356 AF069357 AF069358 AF069359  
AF069360 AF069362 AF069363 AF069364 AF069365 AF069366  
AF069367 AF069368 AF069369 AF069370 AF069371 AF069372 AF069373  
AF069374 AF069375 AF069376 AF069377 AF069378  
AF069333.2 GI:15628184

Homosapiens (human)  
Homosapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 138376)  
Killian, J.K. and Jirtle, R.L.  
Genomic structure of the human M6P/IGF2 receptor  
Mamm. Genome 10 (1), 74-77 (1999)  
99111400  
PUBMED  
9892739  
2 (bases 1 to 138376)  
Killian, J.K. and Jirtle, R.L.  
Allelic variants human M6P/IGF2R

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE









```

variation      /replace="g"
14292
/genes="IGF2R"
/frequency="0.02"
/replace="t"
14341
/genes="IGF2R"
/frequency="0.01"
/replace="t"
14462
/genes="IGF2R"
/frequency="0.04"
/replace="a"
14616
/genes="IGF2R"
/frequency="0.01"
/replace="a"
14655
/genes="IGF2R"
/frequency="0.24"
/replace="a"
14748
/genes="IGF2R"
/frequency="0.01"
/replace="t"
14887
/genes="IGF2R"
/frequency="0.21"
/replace="t"
14893..23111
/genes="IGF2R"
/notes="Region not scanned for variation"
misc_feature

```

```

Query Match      55.0%; Score 26.4; DB 9; Length 139947;
Best Local Similarity 75.0%; Pred.No.19;
Matches 33; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY      1 AGCTCATCCATCTCTGCCAGCCCTCTGCCATCTCTGCTCACTTTC 44
      |||||
Db      44630 AGGTGCTGCATCTCTGTGCTTCTGCCATGACTGTAAGTTTC 44587

```

Search completed: April 28, 2005, 14:18:23  
Job time : 234.469 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 28, 2005, 01:49:33 ; Search time 31.5071 Seconds  
(without alignments)  
9018.523 Million cell updates/sec

Title: US-10-075-846-3\_COPY\_574\_621

Perfect score: 48

Sequence: 1 agctcatcactactgcag.....catctctgtcactttcagtt 48

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq 16Dec04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004as:\*  
13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	100.0	1640	6	AAL49660 Human HGR
2	44	91.7	39796	3	AAC61681 Nucleotid
3	26.4	55.0	339	5	AAS85276 DNA encod
C 4	26	54.2	2848	4	AAS05331 Human alp
C 5	26	54.2	3135	4	AAS05316 Human alp
C 6	26	54.2	3322	4	AAS05322 Human alp
C 7	26	54.2	3558	4	AAS05317 Human alp
C 8	26	54.2	3745	4	AAS05323 Human alp
C 9	26	54.2	73063	12	ADQ97727 Human can
10	24	50.0	35100	2	AAV20441 Human c-f
11	24	50.0	38258	6	AAS98633 DNA encod
12	23.8	49.6	883	10	ADE11700 Human sec
13	23.8	49.6	49380	4	ABL11838 Drosophil
14	23.8	49.6	110000	12	ADQ59443 Human can
C 15	23.6	49.2	418	9	ACH39609 Human foe
16	23.6	49.2	33403	11	ACH44824 Mouse gen
17	23.6	49.2	46338	13	ABD33451 Murine ca
C 18	23.6	49.2	151858	13	ABD33489 Murine ca
19	23.4	48.8	423	4	Aak79953 Human imm
20	23.4	48.8	423	4	Aak79950 Human imm

C 21	23.4	48.8	2069	11	ADM01492 Human cDN
C 22	23.4	48.8	3717	13	ADR08147 Full leng
23	23.4	48.8	67832	9	ADA02801 Mouse itk
24	23.4	48.8	67832	10	ADB72539 Mouse itk
25	23.4	48.8	67832	10	ADC85281 Mouse itk
26	23.4	48.8	67832	12	ADM74396 Murine ca
27	23.2	48.3	454	9	ACH35812 Human end
C 28	23.2	48.3	2964	13	ACN37768 Tumour-as
29	23.2	48.3	5681	12	ADO19507 Human PRO
30	23.2	48.3	67093	13	ABD32671 Mouse can
C 31	23	47.9	717	13	ADS57091 Bacterial
C 32	23	47.9	3678	6	ABS66149 DNA encod
33	23	47.9	10386	4	ABL20052 Drosophil
34	23	47.9	16095	4	ABL04874 Drosophil
C 35	23	47.9	63886	11	ACN44158 Human gen
C 36	23	47.9	66315	12	ADQ97351 Human can
C 37	22.8	47.5	1404	6	ABQ76478 S. cerevi
38	22.8	47.5	50013	11	ACN44708 Mouse gen
39	22.8	47.5	114931	11	ACN44422 Human gen
40	22.6	47.1	3876	6	ADI28076 ECMCAD ge
41	22.6	47.1	4667	6	ABA95199 Human car
42	22.6	47.1	6775	4	AAI58409 Human pol
43	22.6	47.1	6775	5	ADQ98619 DNA encod
44	22.6	47.1	6775	9	ADB48379 Novel hum
45	22.6	47.1	7008	4	AAI60195 Human pol

ALIGNMENTS

RESULT 1  
AAL49660  
ID AAL49660 standard; cDNA; 1640 BP.  
XX  
AC AAL49660;  
XX  
DT 27-NOV-2002 (first entry)  
XX  
DE Human HGRA4 splice variant coding sequence SEQ ID NO: 3.  
XX  
KW Human; glycine receptor alpha subunit 4; HGRA4; HGRA4sv; splice variant;  
KW cardiovascular disorder; reproductive disorder; neural disorder;  
KW cardiac; antiarrhythmic; antianginal; antiarrheic; antiulcer;  
KW nootropic; neuroprotective; antibacterial; virucide; protozoacide;  
KW nervous system disorder; gastrointestinal disorder; gene therapy;  
KW infection; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200266606-A2.  
XX  
PD 29-AUG-2002.  
XX  
PF 13-FEB-2002; 2002WO-US004329.  
XX  
PR 16-FEB-2001; 2001US-0269535P.  
XX  
PA (BRIM ) BRISTOL-MYERS SQUIBB CO.  
XX  
PI Feder J, Lee L, Chen J, Jackson DG, Ramanathan C, Siemers N;  
PI Chang H;  
XX  
XX WPI; 2002-674925/72.  
DR P-PSDB; AAO19187.  
XX  
XX New isolated nucleic acid molecules encoding human glycine receptor A4  
PT (HGRA4) polypeptides, useful for preventing, treating and ameliorating  
PT conditions, e.g. neural or gastrointestinal disorders.  
XX  
PS Claim 1; Fig 2; 349pp; English.  
XX  
CC The present invention provides the protein and coding sequences of the  
CC human glycine receptor alpha 4 (HGRA4) and its splice variant HGRA4sv.

CC The sequences can be used in the treatment of neural disorders,  
CC gastrointestinal disorders, disorders related to hyper glycine receptor  
CC activity, cardiovascular disorders, reproductive disorders, or bacterial,  
CC viral and parasitic infections. The present sequence is a coding sequence  
CC of the invention

XX  
SQ Sequence 1640 BP; 381 A; 456 C; 384 G; 419 T; 0 U; 0 Other;  
Query Match 100.0%; Score 48; DB 6; Length 1640;  
Best Local Similarity 100.0%; Pred. No. 1.1e-07; Mismatches 0; Indels 0; Gaps 0;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGTCATCCATCTCTGCAGCCCTCTGCCATCTCTCTCATTTCAGTT 48  
|||||  
DB 574 AGTCATCCATCTCTGCAGCCCTCTGCCATCTCTCTCATTTCAGTT 621  
|||||

RESULT 2  
AAC61681  
ID AAC61681 standard; DNA; 39796 BP.

XX AAC61681;

XX 19-FEB-2001 (first entry)

XX Nucleotide sequence of the human ataxia gene.

XX Human; ataxia; gene therapy; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT exon 29850..29921

FT intron 29922..33025

FT exon 33026..33155

FT intron 33156..33444

FT exon 33445..33514

FT intron 33513..33751

FT exon 33752..33975

FT intron 33976..34114

FT exon 34115..34195

FT intron 34196..35759

FT exon 35760..35901

FT intron 35902..38781

FT exon 38782..38996

FT intron 38997..40000

FT exon 40001..40000

FT intron 40001..40000

FT exon 40001..40000

FT intron 40001..40000

FT exon 40001..40000

FT intron 40001..40000

FT exon 40001..40000

FT intron 40001..40000

FT exon 40001..40000

FT intron 40001..40000

FT exon 40001..40000

FT intron 40001..40000

FT exon 40001..40000

FT intron 40001..40000

FT exon 40001..40000

FT intron 40001..40000

FT exon 40001..40000

FT intron 40001..40000

FT exon 40001..40000

FT intron 40001..40000

FT exon 40001..40000

FT intron 40001..40000

FT exon 40001..40000

FT intron 40001..40000

FT exon 40001..40000

FT intron 40001..40000

FT exon 40001..40000

PT Novel nucleic acid sequence encoding human ataxia protein for screening  
PT compounds useful for treating disorders relating to mutations in ataxia  
PT gene.

XX Claim 6; Page 22-44; 47pp; English.

XX The present sequence represents the human ataxia gene. The ataxia protein  
CC and polynucleotides are useful for diagnosing and treating disorders  
CC related to ataxia. Ataxia gene sequences are useful in gene therapy, and  
CC as diagnostic tools or reagents for identifying and characterizing  
CC genetic defect involved in the disorders and diseases related to ataxia  
XX

SQ Sequence 39796 BP; 10630 A; 9337 C; 9355 G; 10474 T; 0 U; 0 Other;

Query Match 91.7%; Score 44; DB 3; Length 39796;

Best Local Similarity 100.0%; Pred. No. 6.3e-06; Mismatches 0; Indels 0; Gaps 0;

Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 CATCCATCTCTGCAGCCCTCTGCCATCTCTCTCATTTCAGTT 48

|||||  
DB 35720 CATCCATCTCTGCAGCCCTCTGCCATCTCTCTCATTTCAGTT 35763  
|||||

RESULT 3  
AAS85276

ID AAS85276 standard; cDNA; 399 BP.

XX AAS85276;

XX 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #21080.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX Food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO2001/75067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

XX 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639162/73.

XX P-PSDB; ABG21089.

XX New isolated polynucleotide and encoded polypeptides, useful in

XX diagnostics, forensics, gene mapping, identification of mutations

XX responsible for genetic disorders or other traits and to assess

XX biodiversity.

XX Claim 1; SEQ ID NO 21080; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)

XX sequences. (I) is useful as hybridisation probes, polymerase chain

XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,

XX and in recombinant production of (II). The polynucleotides are also used

XX in diagnostics as expressed sequence tags for identifying expressed

XX genes. (I) is useful in gene therapy techniques to restore normal

XX activity of (II) or to treat disease states involving (II). (II) is

XX useful for generating antibodies against it, detecting or quantitating a

XX polypeptide in tissue, as molecular weight markers and as a food

XX supplement. (II) and its binding partners are useful in medical imaging

XX of sites expressing (II). (I) and (II) are useful for treating disorders

XX involving aberrant protein expression or biological activity. The

CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
 CC coding sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 399 BP; 89 A; 110 C; 98 G; 102 T; 0 U; 0 Other;  
 Query Match 55.0%; Score 26.4; DB 5; Length 399;  
 Best Local Similarity 75.0%; Pred. No. 7.2;  
 Matches 33; Conservative 0; Mismatches 0; Gaps 0;  
 QY 1 AGCTCATCCATCTCTGCAGCCCTCTGCCATCTCTGTCACCTTC 44  
 Db 35 AGGTGCTCGATATCTGTGCTTCTGCCATCTCTGCCATGACTGTAGTTTC 78  
 RESULT 4  
 AAS05331/c  
 ID AAS05331 standard; DNA; 2848 BP.  
 XX  
 AC AAS05331;  
 XX  
 DT 07-SEP-2001 (first entry)  
 XX  
 DE Human alpha-1,3 galactosyltransferase pseudogene genomic sequence #8.  
 XX  
 KW Human; alpha-1,3 galactosyltransferase promoter; livestock;  
 KW recombinant expression cassette; gene expression; transgenic animal;  
 KW growth hormone; xenotransplantation; transgenic cell; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Intron 1..65  
 FT /\*tag= a  
 FT misc\_feature 66..2676  
 FT /\*tag= b  
 FT /\*note= "Untranslated exon 9"  
 FT misc\_feature 2677..2848  
 FT /\*tag= c  
 FT /\*note= "inter-gene sequence"  
 XX  
 PN WO200130992-A2.  
 XX  
 PD 03-MAY-2001.  
 XX  
 PF 20-OCT-2000; 2000WO-US029139.  
 XX  
 PR 22-OCT-1999; 99US-0161092P.  
 PR 25-AUG-2000; 2000US-0227951P.  
 XX  
 PA (UYPI-) UNIV PITTSBURGH.  
 XX  
 PI Koike C;  
 XX  
 DR WPI; 2001-300505/31.  
 XX  
 PS New recombinant expression cassette useful in xenotransplantation.  
 XX  
 PS Claim 16; Page 106-109; 138pp; English.  
 XX  
 CC The present sequence for human alpha-1,3 galactosyltransferase (GT)  
 CC pseudogene genomic sequence #8 is 1 of 44 alpha-1,3 galactosyltransferase  
 CC gene sequences (AAS05290-AAS05333) from different animals. The invention  
 CC provides a novel recombinant expression cassette comprising an alpha-1,3  
 CC GT promoter operably linked to a polynucleotide for expression. The  
 CC cassettes of the invention can be used to express foreign genes or to  
 CC disrupt the native alpha-1,3 GT genomic sequence. The invention provides  
 CC transgenic animals and methods for their production. Transgenic mice in  
 CC which a reporter gene is operably linked to the alpha-1,3 GT promoter can  
 CC be used for assessing promoter activity and specificity. Transgenic  
 CC livestock having the novel expression cassette in which a growth hormone  
 CC is expressed under the control of the alpha-1,3 GT promoter can be

CC transgenic animals and methods for their production. Transgenic mice in  
 CC which a reporter gene is operably linked to the alpha-1,3 GT promoter can  
 CC be used for assessing promoter activity and specificity. Transgenic  
 CC livestock having the novel expression cassette in which a growth hormone  
 CC is expressed under the control of the alpha-1,3 GT promoter can be  
 CC matured or grown better than commonly employed strains. Tissue obtained  
 CC from transgenic animals can be implanted into a host providing a method  
 CC of xenotransplantation from a transgenic animal. The invention also  
 CC provides a transgenic organ consisting essentially of transgenic cells  
 CC engineered, for use in transplantation. The methods of the invention  
 CC facilitate xenotransplantation between species, particularly between  
 CC species exhibiting differential expression of the gal-alpha-gal epitope.  
 CC The invention also describes primers (AAS05334-AAS05385) used to isolate  
 CC the alpha-1,3 GT sequences  
 XX  
 SQ Sequence 2848 BP; 854 A; 558 C; 605 G; 826 T; 0 U; 5 Other;  
 Query Match 54.2%; Score 26; DB 4; Length 2848;  
 Best Local Similarity 76.2%; Pred. No. 15;  
 Matches 32; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
 QY 3 CTCATCCATCTCTGCAGCCCTCTGCCATCTCTGTCACCTTC 44  
 Db 1727 CTCCTCAATCCACTGCAGGCCCTTGCTTCTCTGACCACTC 1686  
 RESULT 5  
 AAS05316/c  
 ID AAS05316 standard; cDNA; 3135 BP.  
 XX  
 AC AAS05316;  
 XX  
 DT 07-SEP-2001 (first entry)  
 XX  
 DE Human alpha-1,3 galactosyltransferase cDNA transcript sequence #1.  
 XX  
 KW Human; alpha-1,3 galactosyltransferase promoter; livestock;  
 KW recombinant expression cassette; gene expression; transgenic animal;  
 KW growth hormone; xenotransplantation; transgenic cell; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200130992-A2.  
 XX  
 PD 03-MAY-2001.  
 XX  
 PF 20-OCT-2000; 2000WO-US029139.  
 XX  
 PR 22-OCT-1999; 99US-0161092P.  
 PR 25-AUG-2000; 2000US-0227951P.  
 XX  
 PA (UYPI-) UNIV PITTSBURGH.  
 XX  
 PI Koike C;  
 XX  
 DR WPI; 2001-300505/31.  
 XX  
 PS New recombinant expression cassette useful in xenotransplantation.  
 XX  
 PS Claim 16; Page 80-83; 138pp; English.  
 XX  
 CC The present sequence for human alpha-1,3 galactosyltransferase (GT) cDNA  
 CC transcript sequence #1 is 1 of 44 alpha-1,3 galactosyltransferase gene  
 CC sequences (AAS05290-AAS05333) from different animals. The invention  
 CC provides a novel recombinant expression cassette comprising an alpha-1,3  
 CC GT promoter operably linked to a polynucleotide for expression. The  
 CC cassettes of the invention can be used to express foreign genes or to  
 CC disrupt the native alpha-1,3 GT genomic sequence. The invention provides  
 CC transgenic animals and methods for their production. Transgenic mice in  
 CC which a reporter gene is operably linked to the alpha-1,3 GT promoter can  
 CC be used for assessing promoter activity and specificity. Transgenic  
 CC livestock having the novel expression cassette in which a growth hormone  
 CC is expressed under the control of the alpha-1,3 GT promoter can be

CC matured or grown better than commonly employed strains. Tissue obtained  
 CC from transgenic animals can be implanted into a host providing a method  
 CC of xenotransplantation from a transgenic animal. The invention also  
 CC provides a transgenic organ consisting essentially of transgenic cells  
 CC engineered, for use in transplantation. The methods of the invention  
 CC facilitate xenotransplantation between species, particularly between  
 CC species exhibiting differential expression of the gal-alpha-gal epitope.  
 CC The invention also describes primers (AAS05334-AAS05385) used to isolate  
 CC the alpha-1,3 GT sequences

XX Sequence 3135 BP; 945-A; 628 C; 704 G; 858 T; 0 U; 0 Other;

Query Match 54.2%; Score 26; DB 4; Length 3135;  
 Best Local Similarity 76.2%; Pred. No. 15;  
 Matches 32; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 3 CTCATCCATCTCTGCAGCCCTTGCATCTCTGTCACCTTC 44  
 |||||  
 Db 2087 CTCTTCATCCACTGCAGGCCCTTGCCTTCTCTGACACCATC 2046

RESULT 6  
 AAS05322/c  
 ID AAS05322 standard; cDNA; 3322 BP.

XX AAS05322;  
 DT 07-SEP-2001 (first entry)

XX Human alpha-1,3 galactosyltransferase cDNA transcript sequence #7.  
 XX Human; alpha-1,3 galactosyltransferase promoter; livestock;  
 KW recombinant expression cassette; gene expression; transgenic animal;  
 KW growth hormone; xenotransplantation; transgenic cell; ss.

XX Homo sapiens.

OS WO200130992-A2.

PN 03-MAY-2001.

PD 20-OCT-2000; 2000WO-US029139.

PF 22-OCT-1999; 99US-0161092P.

PR 25-AUG-2000; 2000US-0227951P.

XX (UYPI-) UNIV PITTSBURGH.

PA Koike C;

XX WPI; 2001-300505/31.

PT New recombinant expression cassette useful in xenotransplantation.

PS Claim 16; Page 94-97; 138pp; English.

XX The present sequence for human alpha-1,3 galactosyltransferase (GT) cDNA  
 CC transcript sequence #7 is 1 of 44 alpha-1,3 galactosyltransferase gene  
 CC sequences (AAS05290-AAS05333) from different animals. The invention  
 CC provides a novel recombinant expression cassette comprising an alpha-1,3  
 CC GT promoter operably linked to a polynucleotide for expression. The  
 CC cassettes of the invention can be used to express foreign genes or to  
 CC disrupt the native alpha-1,3 GT genomic sequence. The invention provides  
 CC transgenic animals and methods for their production. Transgenic mice in  
 CC which a reporter gene is operably linked to the alpha-1,3 GT promoter can  
 CC be used for assessing promoter activity and specificity. Transgenic  
 CC livestock having the novel expression cassette in which a growth hormone  
 CC is expressed under the control of the alpha-1,3 GT promoter can be  
 CC matured or grown better than commonly employed strains. Tissue obtained  
 CC from transgenic animals can be implanted into a host providing a method  
 CC of xenotransplantation from a transgenic animal. The invention also  
 CC provides a transgenic organ consisting essentially of transgenic cells  
 CC engineered, for use in transplantation. The methods of the invention

CC facilitate xenotransplantation between species, particularly between  
 CC species exhibiting differential expression of the gal-alpha-gal epitope.  
 CC The invention also describes primers (AAS05334-AAS05385) used to isolate  
 CC the alpha-1,3 GT sequences

XX Sequence 3322 BP; 991 A; 670 C; 756 G; 905 T; 0 U; 0 Other;

Query Match 54.2%; Score 26; DB 4; Length 3322;  
 Best Local Similarity 76.2%; Pred. No. 16;  
 Matches 32; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 3 CTCATCCATCTCTGCAGCCCTTGCATCTCTGTCACCTTC 44  
 |||||  
 Db 2274 CTCTTCATCCACTGCAGGCCCTTGCCTTCTCTGACACCATC 2233

RESULT 7  
 AAS05317/c  
 ID AAS05317 standard; cDNA; 3558 BP.

XX AAS05317;

XX 07-SEP-2001 (first entry)

XX Human alpha-1,3 galactosyltransferase cDNA transcript sequence #2.

XX Human; alpha-1,3 galactosyltransferase promoter; livestock;

KW recombinant expression cassette; gene expression; transgenic animal;

KW growth hormone; xenotransplantation; transgenic cell; ss.

XX Homo sapiens.

OS WO200130992-A2.

PN 03-MAY-2001.

PD 20-OCT-2000; 2000WO-US029139.

PF 22-OCT-1999; 99US-0161092P.

PR 25-AUG-2000; 2000US-0227951P.

XX (UYPI-) UNIV PITTSBURGH.

PA Koike C;

XX WPI; 2001-300505/31.

PT New recombinant expression cassette useful in xenotransplantation.

PS Claim 16; Page 83-86; 138pp; English.

XX The present sequence for human alpha-1,3 galactosyltransferase (GT) cDNA  
 CC transcript sequence #2 is 1 of 44 alpha-1,3 galactosyltransferase gene  
 CC sequences (AAS05290-AAS05333) from different animals. The invention  
 CC provides a novel recombinant expression cassette comprising an alpha-1,3  
 CC GT promoter operably linked to a polynucleotide for expression. The  
 CC cassettes of the invention can be used to express foreign genes or to  
 CC disrupt the native alpha-1,3 GT genomic sequence. The invention provides  
 CC transgenic animals and methods for their production. Transgenic mice in  
 CC which a reporter gene is operably linked to the alpha-1,3 GT promoter can  
 CC be used for assessing promoter activity and specificity. Transgenic  
 CC livestock having the novel expression cassette in which a growth hormone  
 CC is expressed under the control of the alpha-1,3 GT promoter can be  
 CC matured or grown better than commonly employed strains. Tissue obtained  
 CC from transgenic animals can be implanted into a host providing a method  
 CC of xenotransplantation from a transgenic animal. The invention also  
 CC provides a transgenic organ consisting essentially of transgenic cells  
 CC engineered, for use in transplantation. The methods of the invention  
 CC facilitate xenotransplantation between species, particularly between  
 CC species exhibiting differential expression of the gal-alpha-gal epitope.  
 CC The invention also describes primers (AAS05334-AAS05385) used to isolate  
 CC the alpha-1,3 GT sequences

SQ Sequence 3558 BP; 1060 A; 722 C; 788 G; 988 T; 0 U; 0 Other;  
Query Match 54.2%; Score 26; DB 4; Length 3558;  
Best Local Similarity 76.2%; Pred. No. 16;  
Matches 32; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 3 CTCATCCATCTCTGAGCCCTCTGCCATCTCTGTCACTTTC 44  
DB 2510 CTCTCAATCCACTGAGCCCTTGGCTTCTCTGACACCATC 2469

RESULT 8  
AAS05323/c  
ID AAS05323 standard; cDNA; 3745 BP.  
XX  
AC AAS05323;  
XX  
DT 07-SEP-2001 (first entry)  
XX  
DE Human alpha-1,3 galactosyltransferase cDNA transcript sequence #8.  
XX  
KW Human; alpha-1,3 galactosyltransferase promoter; livestock;  
KW recombinant expression cassette; gene expression; transgenic animal;  
KW growth hormone; xenotransplantation; transgenic cell; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200130992-A2.  
XX  
PD 03-MAY-2001.  
XX  
PF 20-OCT-2000; 2000WO-US029139.  
XX  
PR 22-OCT-1999; 99US-0161092P.  
PR 25-AUG-2000; 2000US-0227951P.  
XX  
PA (UYPI-) UNIV PITTSBURGH.  
XX  
PI Koike C;  
XX  
DR WPI; 2001-300505/31.  
XX  
PT New recombinant expression cassette useful in xenotransplantation.  
XX  
PS Claim 16; Page 97-100; 138pp; English.  
XX  
CC The present sequence for human alpha-1,3 galactosyltransferase (GT)-cDNA  
CC transcript sequence #8 is 1 of 44 alpha-1,3 galactosyltransferase gene  
CC sequences (AAS05230-AAS05333) from different animals. The invention  
CC provides a novel recombinant expression cassette comprising an alpha-1,3  
CC GT promoter operably linked to a polynucleotide for expression. The  
CC cassettes of the invention can be used to express foreign genes or to  
CC disrupt the native alpha-1,3 GT genomic sequence. The invention provides  
CC transgenic animals and methods for their production. Transgenic mice in  
CC which a reporter gene is operably linked to the alpha-1,3 GT promoter can  
CC be used for assessing promoter activity and specificity. Transgenic  
CC livestock having the novel expression cassette in which a growth hormone  
CC is expressed under the control of the alpha-1,3 GT promoter can be  
CC matured or grown better than commonly employed strains. Tissue obtained  
CC from transgenic animals can be implanted into a host providing a method  
CC of xenotransplantation from a transgenic animal. The invention also  
CC provides a transgenic organ consisting essentially of transgenic cells  
CC engineered, for use in transplantation. The methods of the invention  
CC facilitate xenotransplantation between species, particularly between  
CC species exhibiting differential expression of the gal-alpha-gal epitope.  
CC The invention also describes primers (AAS05334-AAS05385) used to isolate  
CC the alpha-1,3 GT sequences  
XX

SQ Sequence 3745 BP; 1106 A; 764 C; 840 G; 1035 T; 0 U; 0 Other;  
Query Match 54.2%; Score 26; DB 4; Length 3745;  
Best Local Similarity 76.2%; Pred. No. 16;  
Matches 32; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 3 CTCATCCATCTCTGAGCCCTCTGCCATCTCTGTCACTTTC 44  
DB 2697 CTCTCAATCCACTGAGCCCTTGGCTTCTCTGACACCATC 2656

RESULT 9  
ADQ97727/c  
ID ADQ97727 standard; DNA; 73063 BP.  
XX  
AC ADQ97727;  
XX  
DT 07-OCT-2004 (first entry)  
XX  
DE Human cancer associated sequence HD10-034, SEQ ID 704.  
XX  
KW Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Human; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO2004060304-A2.  
XX  
PD 22-JUL-2004.  
XX  
PF 22-DEC-2003; 2003WO-US041389.  
XX  
PR 27-DEC-2002; 2002US-00330773.  
XX  
PA (SAGR-) SAGRES DISCOVERY INC.  
XX  
PI Morris DW, Malandro MS;  
XX  
DR WPI; 2004-543781/52.  
XX  
PT New isolated cancer associated nucleic acids comprising at least 10  
PT contiguous nucleotides, useful for diagnosing, preventing and/or treating  
PT cancers such as leukemia and lymphoma.  
XX  
PS Claim 1; SEQ ID NO 704; 199pp; English.  
XX  
CC The present invention relates to cancer associated sequences (ADQ97025-  
CC ADQ98004). The sequences are useful for the diagnosis, prevention and/or  
CC treatment of cancer, such as leukemia and lymphoma. Note: The sequence  
CC data for this patent did not form part of the printed specification, but  
CC was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 73063 BP; 19276 A; 16563 C; 17075 G; 19769 T; 0 U; 380 Other;  
Query Match 54.2%; Score 26; DB 12; Length 73063;  
Best Local Similarity 76.2%; Pred. No. 30;  
Matches 32; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 3 CTCATCCATCTCTGAGCCCTCTGCCATCTCTGTCACTTTC 44  
DB 64035 CTCTCAATCCACTGAGCCCTTGGCTTCTCTGACACCATC 63994

RESULT 10  
AAV20441  
ID AAV20441 standard; DNA; 35100 BP.  
XX  
AC AAV20441;  
XX  
DT 17-JUN-1998 (first entry)  
XX  
DE Human c-fms oncogene.  
XX  
KW Human; oncogene; proto-oncogene; neoplastic disease; anticancer; cancer;  
KW antisense oligonucleotide; c-fms; ds.  
XX  
OS Homo sapiens.  
XX









PT cancer (e.g. lymphoma) or for screening agents that may be used for  
PT treating or preventing cancer.  
XX  
XX  
XX Disclosure; SEQ ID NO 79; 143pp; English.  
XX  
XX The present invention describes human cancer-associated (CA) nucleotide  
CC sequences (1). Also described: (1) an expression vector comprising (1);  
CC (2) a host cell comprising (1) or the expression vector; (3) a microarray  
CC for detecting a CA nucleic acid; (4) an isolated polypeptide encoded  
CC within an open reading frame of a CA sequence; (5) an isolated antibody,  
CC or its antigen binding fragment, that binds to the above polypeptide; (6)  
CC a hybridoma that produces the monoclonal antibody described above; (7) a  
CC pharmaceutical composition comprising the antibody and a pharmaceutical  
CC excipient; (8) a kit for detecting or diagnosing cancer cells, comprising  
CC the above (monoclonal) antibody or polynucleotide that selectively  
CC hybridizes to any of the polynucleotide sequences mentioned above; (9)  
CC methods for diagnosing cancer or for detecting the presence or absence of  
CC cancer cells in an individual; (10) a method for inhibiting growth of  
CC cancer cells in an individual; (11) a method for delivering a therapeutic  
CC agent to cancer cells in an individual; (12) an electronic library  
CC comprising the polynucleotide or polypeptide, or their fragments,  
CC mentioned above; (13) a method of screening for anticancer activity; (14)  
CC methods for detecting cancer associated with expression of a polypeptide  
CC or the presence of the antibody in a test cell or serum sample; (15) a  
CC method for screening for a bioactive agent capable of modulating the  
CC activity of a CA protein encoded by the above nucleic acid molecule; and  
CC (16) a method for treating cancers. (1) has cytostatic activity, and can  
CC be used in gene therapy, in vaccines, as a tyrosine kinase antagonist,  
CC and as a G-protein coupled receptor antagonist. The compositions and  
CC methods of the present invention can be used for diagnosing, preventing  
CC and treating cancer, especially lymphomas. They may also be used in  
CC screening for agents that may be used for treating or preventing cancer.  
CC The present sequence represents a human CA gene sequence, which is given  
CC in the exemplification of the present invention. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 352938 BP; 94852 A; 72623 C; 75909 G; 104334 T; 0 U; 5220 Other;  
SQ  
Query Match 49.6%; Score 23.8; DB 12; Length 110000;  
Best Local Similarity 72.1%; Pred. No. 2.1e+02;  
Matches 31; Conservative 0; Mismatches 12; Indels 0; Gaps 0;  
QY 6 ATCCATCTCTGAGCCCTCTGCCATCTCTGTCACCTTCAGTT 48  
Db 46646 ATGCATCTCTGAGACCACCTCCAGCCCTGTGCCATTCACCT 46688  
RESULT 15  
ACH39609/c  
ID ACH39609 standard; cDNA; 418 BP.  
XX  
XX ACH39609;  
XX  
XX 13-OCT-2003 (first entry)  
XX  
XX Human foetal brain cDNA #976.  
XX  
XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;  
KW Genome mapping; biodiversity; genetic disorder.  
XX  
XX Homo sapiens.  
XX  
XX US2003073623-A1.  
XX  
XX 17-APR-2003.  
XX  
XX 30-JUL-2001; 2001US-00918995.  
XX  
XX 30-JUL-2001; 2001US-00918995.  
XX  
XX (DRMA/) DRMANAC R T.  
PA  
(LABA/) LABAT I.  
(STAC/) STACHE-CRAIN B.  
PA (DICK/) DICKSON M C.  
PA (JONE/) JONES L W.  
XX  
XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;  
XX WPI; 2003-615964/58.  
XX  
XX New polynucleotide sequences obtained from various cDNA libraries, useful  
PT as hybridization probes, as oligomers for PCR, for chromosome and gene  
PT mapping, in the recombinant production of protein, or in generating  
PT antisense DNA or RNA.  
XX  
XX Claim 1; SEQ ID NO 26821; 44pp; English.  
XX  
XX The invention relates to an isolated polynucleotide comprising any one of  
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was  
CC determined by the technique of SBH (sequencing by hybridisation). Also  
CC included is a purified polypeptide comprising a sequence corresponding to  
CC a reading frame of the novel polynucleotide. The nucleic acid sequences  
CC are useful in diagnostics as expressed sequence tags (EST) for  
CC identifying expressed genes or for physical mapping of the human genome,  
CC in forensics, in assessing biodiversity, or in identifying mutations  
CC responsible for genetic disorders and other traits. The nucleotide  
CC sequences are also useful as hybridisation probes, as oligomers for PCR,  
CC for chromosome and gene mapping, in the recombinant production of  
CC protein, or in generating antisense DNA or RNA. The purified polypeptide  
CC is useful for generating antibodies specific for it. The present sequence  
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?DocID=20030073623  
XX  
XX Sequence 418 BP; 145 A; 52 C; 129 G; 91 T; 0 U; 1 Other;  
SQ  
Query Match 49.2%; Score 23.6; DB 9; Length 418;  
Best Local Similarity 86.7%; Pred. No. 78;  
Matches 26; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 5 CATCCATCTCTGCGAGCCCTCTGCCATCTC 34  
Db 244 CCTCCAGACTCTGCATCCCTCTGCTTCTC 215  
Search completed: April 28, 2005, 12:06:05  
Job time : 38.5071 secs

***This Page Blank (usp10)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 28, 2005, 11:47:18 ; Search time 9.8673 Seconds  
(without alignments)  
7959.755 Million cells updates/sec

Title: US-10-075-846-3\_COPY\_574\_621

Perfect score: 48  
Sequence: 1 agctcatcactactctgcag.....catctctgtcactttcagtt 48

Scoring table: IDENTITY\_NUC

Gapop 10\_0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.\*

- 1: /cgn2\_6/prodata/1/ina/5A COMB.seq.\*
- 2: /cgn2\_6/prodata/1/ina/5B COMB.seq.\*
- 3: /cgn2\_6/prodata/1/ina/6A COMB.seq.\*
- 4: /cgn2\_6/prodata/1/ina/6B COMB.seq.\*
- 5: /cgn2\_6/prodata/1/ina/PCTUS COMB.seq.\*
- 6: /cgn2\_6/prodata/1/ina/backfile1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	24	50.0	35100	1	US-08-306-691B-19
2	24	50.0	35100	5	PCT-US93-06251-19
C 3	23.2	48.3	601	4	US-09-949-016-191370
C 4	23.2	48.3	42276	4	US-09-949-016-17218
5	23	47.9	601	4	US-09-949-016-135808
6	23	47.9	601	4	US-09-949-016-135925
7	23	47.9	601	4	US-09-949-016-136042
C 8	23	47.9	193555	4	US-09-949-016-15553
C 9	23	47.9	193555	4	US-09-949-016-15554
C 10	23	47.9	193555	4	US-09-949-016-15555
11	23	47.9	222452	4	US-09-949-016-12968
12	22.6	47.1	6775	4	US-09-620-312D-289
C 13	22.4	46.7	304533	4	US-09-949-016-15371
C 14	22.4	46.7	304533	4	US-09-949-016-15372
C 15	22.2	46.2	78491	4	US-09-949-016-15132
C 16	22	45.8	601	4	US-09-949-016-15132
17	22	45.8	825	1	US-08-248-474-109
18	22	45.8	825	3	US-08-756-849-109
C 19	22	45.8	44789	4	US-09-949-016-13909
C 20	21.8	45.4	601	4	US-09-949-016-127237
C 21	21.8	45.4	5761	3	US-09-323-472A-1
C 22	21.8	45.4	5761	3	US-09-323-472A-3
C 23	21.8	45.4	5761	3	US-09-323-472A-11
C 24	21.8	45.4	5761	4	US-09-585-077C-1
C 25	21.8	45.4	5761	4	US-09-585-077C-3
C 26	21.8	45.4	5761	4	US-09-585-077C-11
C 27	21.8	45.4	5761	4	US-09-585-077C-13

C 28	21.8	45.4	56939	4	US-09-949-016-13613	Sequence 13613, A
C 29	21.8	45.4	65561	4	US-09-949-016-13665	Sequence 15365, A
C 30	21.8	45.4	76281	4	US-09-949-016-12708	Sequence 12708, A
C 31	21.8	45.4	150780	4	US-09-949-016-14711	Sequence 14711, A
C 32	21.6	45.0	423	4	US-09-513-999C-28242	Sequence 28242, A
C 33	21.6	45.0	432	4	US-09-513-999C-1913	Sequence 1913, Appl
C 34	21.6	45.0	447	1	US-08-253-155A-22	Sequence 22, Appl
C 35	21.6	45.0	601	4	US-09-949-016-25301	Sequence 25301, A
C 36	21.6	45.0	601	4	US-09-949-016-25302	Sequence 25302, A
C 37	21.6	45.0	601	4	US-09-949-016-81829	Sequence 81829, A
C 38	21.6	45.0	601	4	US-09-949-016-81830	Sequence 81830, A
C 39	21.6	45.0	601	4	US-09-949-016-81831	Sequence 81831, A
C 40	21.6	45.0	601	4	US-09-949-016-81832	Sequence 81832, A
C 41	21.6	45.0	601	4	US-09-949-016-14260	Sequence 124260, A
C 42	21.6	45.0	601	4	US-09-949-016-14261	Sequence 124261, A
C 43	21.6	45.0	601	4	US-09-949-016-144834	Sequence 144834, A
C 44	21.6	45.0	601	4	US-09-949-016-144835	Sequence 144835, A
C 45	21.6	45.0	601	4	US-09-949-016-144836	Sequence 144836, A

ALIGNMENTS

RESULT 1  
US-08-306-691B-19  
; Sequence 19, Application US/08306691B  
; Patent No. 5734039  
; GENERAL INFORMATION:  
; APPLICANT: Calabretta, Bruno  
; APPLICANT: Skorski, Tomasz  
; TITLE OF INVENTION: ANTISENSE  
; TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES  
; NUMBER OF SEQUENCES: 55  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seidel, Gonda, Lavorigna & Monaco, P.C.  
; STREET: Two Penn Center, Suite 1800  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: U.S.A.  
; ZIP: 19102  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/306,691B  
; FILING DATE: September 15, 1994  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Monaco, Daniel A.  
; REGISTRATION NUMBER: 30,480  
; REFERENCE/DOCKET NUMBER: 8321-8  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-8383  
; TELEFAX: (215) 568-5549  
; TELEX: No. 5734039e  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 35100 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
US-08-306-691B-19

Query Match 50.0%; Score 24; DB 1; Length 35100;  
Best Local Similarity 75.0%; Pred. No. 44;  
Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 2 GCTCATCTCACTCTCGACCCCTCTGCCATCTCTGTCACT 41





Query Match 47.9%; Score 23; DB 4; Length 601;  
Best Local Similarity 68.1%; Pred. No. 41;  
Matches 32; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Query Match 47.9%; Score 23; DB 4; Length 193555;

Best Local Similarity 68.1%; Pred. No. 1.5e+02;  
Matches 32; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 AGCTCATCATCTGCGAGCCCTCTGCCATCTCTGTCACTTTCACT 47  
DB 73294 AGCTCTGCAGCCTCTGCACCTGTGCTGCCACCACTGCCACTGTGAAT 73248

RESULT 10  
US-09-949-016-15555/c  
; Sequence 15555, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15555  
; LENGTH: 193555  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)-(193555)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-15555

Query Match 47.9%; Score 23; DB 4; Length 193555;  
Best Local Similarity 68.1%; Pred. No. 1.5e+02;  
Matches 32; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 AGCTCATCATCTGCGAGCCCTCTGCCATCTCTGTCACTTTCACT 47  
DB 73294 AGCTCTGCAGCCTCTGCACCTGTGCTGCCACCACTGCCACTGTGAAT 73248

RESULT 11  
US-09-949-016-12968  
; Sequence 12968, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12968  
; LENGTH: 222452  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-12968

Query Match 47.9%; Score 23; DB 4; Length 222452;  
Best Local Similarity 68.1%; Pred. No. 1.6e+02;

Matches 32; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 2 GCTCATCATCTGCGAGCCCTCTGCCATCTCTGTCACTTTCACTT 48  
DB 67657 GCGCTCCCTCCACTCTGCGAGCCCTCTGCCACCACTGTGAGATT 67703

RESULT 12  
US-09-620-312D-289  
; Sequence 289, Application US/09620312D  
; Patent No. 6569682  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Zhang, Jie  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Chen, Rui-bong  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Yang, Yonghong  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Zhou, Ping  
; APPLICANT: Ma, Yunging  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: John Tillinghast  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and  
; TITLE OF INVENTION: Polypeptides  
; FILE REFERENCE: 784CIP2B  
; CURRENT APPLICATION NUMBER: US/09/620,312D  
; CURRENT FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: 09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 1105  
; SOFTWARE: pt\_FL\_genes Version 1.0  
; SEQ ID NO 289  
; LENGTH: 6775  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (197)..(4471)  
US-09-620-312D-289

Query Match 47.1%; Score 22.6; DB 4; Length 6775;  
Best Local Similarity 68.9%; Pred. No. 1e+02;  
Matches 31; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 AGCTCATCATCTGCGAGCCCTCTGCCATCTCTGTCACTTTCA 45  
DB 4032 AACCCATCCATTCAGCAGCCAGGTACAGCCTCTGCCACCTCCA 4076

RESULT 13  
US-09-949-016-15371/c  
; Sequence 15371, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03

US-09-949-016-15132/c  
 ; Sequence 15132, Application US/09949016  
 ; Patent No. 6812339  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig et al.  
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
 ; FILE REFERENCE: CL001307  
 ; CURRENT APPLICATION NUMBER: US/09/949,016  
 ; CURRENT FILING DATE: 2000-04-14  
 ; PRIOR APPLICATION NUMBER: 60/241,755



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 28, 2005, 08:01:55 ; Search time 89.6588 Seconds  
(without alignments)  
3259.123 Million cell updates/sec

Title: US-10-075-846-3\_COPY\_574\_621

Perfect score: 48

Sequence: 1 agctcatcactactgcag.....catctctgtcactttcagtt 48

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5642217 seqs, 3043843248 residues

Total number of hits satisfying chosen parameters: 1128434

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq\*
- 18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq\*
- 19: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq\*
- 20: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq\*
- 21: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq\*
- 22: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	100.0	1640	14	US-10-075-846-3 Sequence 3, Appli
C 2	26.4	55.0	201	19	US-10-741-600-36664 Sequence 36664, A
C 3	26.4	55.0	149382	19	US-10-741-600-17661 Sequence 17661, A
C 4	26	54.2	2848	17	US-10-125-994A-42 Sequence 42, Appl
C 5	26	54.2	3135	17	US-10-125-994A-27 Sequence 27, Appl
C 6	26	54.2	3322	17	US-10-125-994A-33 Sequence 33, Appl
C 7	26	54.2	3558	17	US-10-125-994A-28 Sequence 28, Appl
C 8	26	54.2	3745	17	US-10-125-994A-34 Sequence 34, Appl
C 9	24.2	50.4	548	18	US-10-437-963-43806 Sequence 43806, A
10	24	50.0	201	19	US-10-741-600-56429 Sequence 56429, A
11	24	50.0	201	19	US-10-741-600-56504 Sequence 56504, A

12	24	50.0	201	19	US-10-741-600-56505	Sequence 56505, A
13	24	50.0	72069	19	US-10-741-600-17775	Sequence 17775, A
14	23.8	49.6	883	10	US-09-892-877-73	Sequence 73, Appl
15	23.8	49.6	883	10	US-09-948-783-72	Sequence 72, Appl
16	23.8	49.6	352938	18	US-10-322-696-79	Sequence 79, Appl
17	23.6	49.2	263	17	US-10-424-599-31798	Sequence 31798, A
18	23.6	49.2	418	10	US-09-918-995-26821	Sequence 26821, A
C 19	23.6	49.2	33403	13	US-10-087-132-1465	Sequence 1465, Ap
20	23.6	49.2	46338	18	US-10-322-281-595	Sequence 595, App
C 21	23.6	49.2	151858	18	US-10-322-281-653	Sequence 653, App
22	23.4	48.8	608	13	US-10-027-632-187595	Sequence 187595,
23	23.4	48.8	608	13	US-10-027-632-187595	Sequence 187595,
C 24	23.4	48.8	2069	17	US-10-108-260A-177	Sequence 177, App
C 25	23.4	48.8	65300	19	US-10-741-600-17776	Sequence 17776, A
26	23.4	48.8	67832	11	US-09-997-722-67	Sequence 67, Appl
C 27	23.2	48.3	413	13	US-10-027-632-288605	Sequence 288605,
C 28	23.2	48.3	413	13	US-10-027-632-288605	Sequence 288605,
29	23.2	48.3	454	10	US-09-918-995-23024	Sequence 23024, A
30	23.2	48.3	67093	18	US-10-417-375-33	Sequence 33, Appl
C 31	23	47.9	655	13	US-10-027-632-12623	Sequence 12623, A
C 32	23	47.9	655	13	US-10-027-632-12624	Sequence 12624, A
C 33	23	47.9	655	13	US-10-027-632-12625	Sequence 12625, A
C 34	23	47.9	655	17	US-10-027-632-12623	Sequence 12623, A
C 35	23	47.9	655	17	US-10-027-632-12623	Sequence 12623, A
C 36	23	47.9	655	17	US-10-027-632-12625	Sequence 12625, A
C 37	23	47.9	717	17	US-10-369-493-32765	Sequence 32765, A
C 38	23	47.9	809	13	US-10-027-632-167429	Sequence 167429,
C 39	23	47.9	809	17	US-10-027-632-167429	Sequence 167429,
C 40	23	47.9	3678	17	US-10-432-443-140	Sequence 140, App
C 41	23	47.9	63686	13	US-10-087-192-466	Sequence 466, App
C 42	22.8	47.5	700	18	US-10-425-115-44389	Sequence 44389, A
C 43	22.8	47.5	1404	18	US-10-451-467A-381	Sequence 381, App
44	22.8	47.5	50013	13	US-10-087-192-1291	Sequence 1291, Ap
45	22.8	47.5	114931	13	US-10-087-192-862	Sequence 862, App

## ALIGNMENTS

### RESULT 1

US-10-075-846-3  
; Sequence 3, Application US/10075846  
; Publication No. US20030032608A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING A NOVEL GLYCINE RECEPTOR ALPHA SUBUNIT THEREOF  
; TITLE OF INVENTION: IN THE GASTROINTESTINAL TRACT, HGRA4, and SPLICE VARIANT THEREOF  
; FILE REFERENCE: D0079 NP  
; CURRENT APPLICATION NUMBER: US/10/075,846  
; CURRENT FILING DATE: 2002-02-13  
; PRIOR APPLICATION NUMBER: US 60/269,535  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 81  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 3  
; LENGTH: 1640  
; TYPE: DNA  
; ORGANISM: homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1293)  
US-10-075-846-3

Query Match 100.0%; Score 48; DB 14; Length 1640;  
Best Local Similarity 100.0%; Pred. No. 3.3e-08;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCTCATCATACTCTGCGAGCCCTCTGCCATCTCTGTCACTTCAGTT 48

Db 574 AGCTCATCATACTCTGCGAGCCCTCTGCCATCTCTGTCACTTCAGTT 621

### RESULT 2

```
US-10-741-600-36664/c
; Sequence 36664, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36664
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-36664

Query Match      55.0%; Score 26.4; DB 19; Length 201;
Best Local Similarity 75.0%; Pred. No. 3.6;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCTCATCCATCTCTGCAGCCCTCTGCCATCTCTGTCACTTTC 44
   |||||
Db 194 AGTGCTCGATCTCTGTTCCTTCTGCCATGACTGTAAATTTC 151

RESULT 3
US-10-741-600-17661/c
; Sequence 17661, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17661
; LENGTH: 149382
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(149382)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-741-600-17661

Query Match      55.0%; Score 26.4; DB 19; Length 149382;
Best Local Similarity 75.0%; Pred. No. 5.1;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCTCATCCATCTCTGCAGCCCTCTGCCATCTCTGTCACTTTC 44
   |||||
Db 49389 AGTGCTCGATCTCTGTTCCTTCTGCCATGACTGTAAATTTC 49346

RESULT 4
US-10-125-994A-42/c
; Sequence 42, Application US/10125994A
; Publication No. US20030203427A1
; GENERAL INFORMATION:
; APPLICANT: Koike, Chihiro
; TITLE OF INVENTION: ALPHA 1,3-GALACTOSYLTRANSFERASE GENE AND PROMOTER
; FILE REFERENCE: 206779
; CURRENT APPLICATION NUMBER: US/10/125,994A
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: PCT/US00/29139
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 60/227,951
; PRIOR FILING DATE: 2000-09-25
; CURRENT FILING DATE: 2002-04-19
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-125-994A-42/c

Query Match      55.0%; Score 26.4; DB 19; Length 149382;
Best Local Similarity 75.0%; Pred. No. 5.1;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCTCATCCATCTCTGCAGCCCTCTGCCATCTCTGTCACTTTC 44
   |||||
Db 49389 AGTGCTCGATCTCTGTTCCTTCTGCCATGACTGTAAATTTC 49346

RESULT 4
US-10-125-994A-42/c
; Sequence 42, Application US/10125994A
; Publication No. US20030203427A1
; GENERAL INFORMATION:
; APPLICANT: Koike, Chihiro
; TITLE OF INVENTION: ALPHA 1,3-GALACTOSYLTRANSFERASE GENE AND PROMOTER
; FILE REFERENCE: 206779
; CURRENT APPLICATION NUMBER: US/10/125,994A
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: PCT/US00/29139
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 60/227,951
; PRIOR FILING DATE: 2000-09-25
; CURRENT FILING DATE: 2002-04-19
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
```

```
US-10-125-994A-42/c
; Sequence 42, Application US/10125994A
; Publication No. US20030203427A1
; GENERAL INFORMATION:
; APPLICANT: Koike, Chihiro
; TITLE OF INVENTION: ALPHA 1,3-GALACTOSYLTRANSFERASE GENE AND PROMOTER
; FILE REFERENCE: 206779
; CURRENT APPLICATION NUMBER: US/10/125,994A
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: PCT/US00/29139
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 60/227,951
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 60/161,092
; PRIOR FILING DATE: 1999-10-22
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 27
; LENGTH: 3135
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-125-994A-27

Query Match      54.2%; Score 26; DB 17; Length 3135;
Best Local Similarity 76.2%; Pred. No. 5.9;
Matches 32; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 3 CTCATCCATCTCTGCAGCCCTCTGCCATCTCTGTCACTTTC 44
   |||||
Db 1727 CTCCTCAATCCACTGCAGGCCCTTTCCTTCTCTGACCATC 1686

RESULT 5
US-10-125-994A-27/c
; Sequence 27, Application US/10125994A
; Publication No. US20030203427A1
; GENERAL INFORMATION:
; APPLICANT: Koike, Chihiro
; TITLE OF INVENTION: ALPHA 1,3-GALACTOSYLTRANSFERASE GENE AND PROMOTER
; FILE REFERENCE: 206779
; CURRENT APPLICATION NUMBER: US/10/125,994A
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: PCT/US00/29139
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 60/227,951
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 60/161,092
; PRIOR FILING DATE: 1999-10-22
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 27
; LENGTH: 3135
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-125-994A-27

Query Match      54.2%; Score 26; DB 17; Length 3135;
Best Local Similarity 76.2%; Pred. No. 5.9;
Matches 32; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 3 CTCATCCATCTCTGCAGCCCTCTGCCATCTCTGTCACTTTC 44
   |||||
Db 1727 CTCCTCAATCCACTGCAGGCCCTTTCCTTCTCTGACCATC 1686

RESULT 6
US-10-125-994A-33/c
; Sequence 33, Application US/10125994A
; Publication No. US20030203427A1
; GENERAL INFORMATION:
; APPLICANT: Koike, Chihiro
; TITLE OF INVENTION: ALPHA 1,3-GALACTOSYLTRANSFERASE GENE AND PROMOTER
; FILE REFERENCE: 206779
; CURRENT APPLICATION NUMBER: US/10/125,994A
; CURRENT FILING DATE: 2002-04-19
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-125-994A-33/c

Query Match      54.2%; Score 26; DB 17; Length 3135;
Best Local Similarity 76.2%; Pred. No. 5.9;
Matches 32; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 3 CTCATCCATCTCTGCAGCCCTCTGCCATCTCTGTCACTTTC 44
   |||||
Db 2087 CTCCTCAATCCACTGCAGGCCCTTTCCTTCTCTGACCATC 2046

RESULT 6
US-10-125-994A-33/c
; Sequence 33, Application US/10125994A
; Publication No. US20030203427A1
; GENERAL INFORMATION:
; APPLICANT: Koike, Chihiro
; TITLE OF INVENTION: ALPHA 1,3-GALACTOSYLTRANSFERASE GENE AND PROMOTER
; FILE REFERENCE: 206779
; CURRENT APPLICATION NUMBER: US/10/125,994A
; CURRENT FILING DATE: 2002-04-19
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
```

; PRIOR APPLICATION NUMBER: PCT/US00/29139  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: US 60/227,951  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US 60/161,092  
; PRIOR FILING DATE: 1999-10-22  
; NUMBER OF SEQ ID NOS: 96  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 33  
; LENGTH: 3322  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-125-994A-33

Query Match 54.2%; Score 26; DB 17; Length 3322;  
Best Local Similarity 76.2%; Pred. No. 5.9;  
Matches 32; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 3 CTCATCCATACTCTGCAGCCCTCTGCCATCTCTGTCACTTTC 44  
DB 2274 CTCTTCAATCCACTGCAGGCCCTTGGCTTCTCTGACACCATC 2233

RESULT 7  
US-10-125-994A-28/c  
; Sequence 28, Application US/10125994A  
; Publication No. US20030203427A1  
; GENERAL INFORMATION:  
; APPLICANT: Koike, Chihiro  
; TITLE OF INVENTION: ALPHA 1,3-GALACTOSYLTRANSFERASE GENE AND PROMOTER  
; FILE REFERENCE: 206779  
; CURRENT APPLICATION NUMBER: US/10/125,994A  
; CURRENT FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: PCT/US00/29139  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: US 60/227,951  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US 60/161,092  
; PRIOR FILING DATE: 1999-10-22  
; NUMBER OF SEQ ID NOS: 96  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 28  
; LENGTH: 3558  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-125-994A-28

Query Match 54.2%; Score 26; DB 17; Length 3558;  
Best Local Similarity 76.2%; Pred. No. 5.9;  
Matches 32; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 3 CTCATCCATACTCTGCAGCCCTCTGCCATCTCTGTCACTTTC 44  
DB 2510 CTCTTCAATCCACTGCAGGCCCTTGGCTTCTCTGACACCATC 2469

RESULT 8  
US-10-125-994A-34/c  
; Sequence 34, Application US/10125994A  
; Publication No. US20030203427A1  
; GENERAL INFORMATION:  
; APPLICANT: Koike, Chihiro  
; TITLE OF INVENTION: ALPHA 1,3-GALACTOSYLTRANSFERASE GENE AND PROMOTER  
; FILE REFERENCE: 206779  
; CURRENT APPLICATION NUMBER: US/10/125,994A  
; CURRENT FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: PCT/US00/29139  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: US 60/227,951  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US 60/161,092  
; PRIOR FILING DATE: 1999-10-22  
; NUMBER OF SEQ ID NOS: 96

; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 34  
; LENGTH: 3745  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-125-994A-34

Query Match 54.2%; Score 26; DB 17; Length 3745;  
Best Local Similarity 76.2%; Pred. No. 5.9;  
Matches 32; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 3 CTCATCCATACTCTGCAGCCCTCTGCCATCTCTGTCACTTTC 44  
DB 2697 CTCTTCAATCCACTGCAGGCCCTTGGCTTCTCTGACACCATC 2656

RESULT 9  
US-10-437-963-43806/c  
; Sequence 43806, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 43806  
; LENGTH: 548  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_46927C.1  
US-10-437-963-43806

Query Match 50.4%; Score 24.2; DB 18; Length 548;  
Best Local Similarity 78.4%; Pred. No. 25;  
Matches 29; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2 GCTCATCCATACTCTGCAGCCCTCTGCCATCTCTGTCTC 38  
DB 68 GCTCTCCATACGCTGGCGCCACACTGCCGTCCCTCTC 32

RESULT 10  
US-10-741-600-56429  
; Sequence 56429, Application US/10741600  
; Publication No. US20050026169A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001499  
; CURRENT APPLICATION NUMBER: US/10/741,600  
; CURRENT FILING DATE: 2003-12-22  
; NUMBER OF SEQ ID NOS: 73997  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 56429  
; LENGTH: 201  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-741-600-56429

Query Match 50.0%; Score 24; DB 19; Length 201;  
Best Local Similarity 75.0%; Pred. No. 28;



```
Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 2 GCTATCCATACCTCGAGCCCTCTGCCATCTCTGTCACT 41
Db 13 GCTATCCCTTCTCTCTCTCTCTCTGCCATCTGTGTCT 52

RESULT 11
US-10-741-600-56504
; Sequence 56504, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; PRIORITY FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 56504
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-56504

Query Match 50.0%; Score 24; DB 19; Length 201;
Best Local Similarity 75.0%; Pred. No. 28;
Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 2 GCTATCCATACCTCGAGCCCTCTGCCATCTCTGTCACT 41
Db 91 GCTATCCCTTCTCTCTCTCTCTCTGCCATCTGTGTCT 130

RESULT 12
US-10-741-600-56505
; Sequence 56505, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; PRIORITY FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 56505
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-56505

Query Match 50.0%; Score 24; DB 19; Length 201;
Best Local Similarity 75.0%; Pred. No. 28;
Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 2 GCTATCCATACCTCGAGCCCTCTGCCATCTCTGTCACT 41
Db 12 GCTATCCCTTCTCTCTCTCTCTCTGCCATCTGTGTCT 51

RESULT 13
US-10-741-600-17775
; Sequence 17775, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
```

```
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17775
; LENGTH: 72069
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-17775

Query Match 50.0%; Score 24; DB 19; Length 72069;
Best Local Similarity 75.0%; Pred. No. 39;
Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 2 GCTATCCATACCTCGAGCCCTCTGCCATCTCTGTCACT 41
Db 58192 GCTATCCCTTCTCTCTCTCTCTCTGCCATCTGTGTCT 58231

RESULT 14
US-09-892-877-73
; Sequence 73, Application US/09892877
; Publication No. US20030077809A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et. al.
; TITLE OF INVENTION: 97 Human secreted proteins
; FILE REFERENCE: P2028P1
; CURRENT APPLICATION NUMBER: US/09/892,877
; PRIORITY FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/437,658
; PRIORITY FILING DATE: EARLIER FILING DATE: 1999-11-10
; NUMBER OF SEQ ID NOS: 461
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 73
; LENGTH: 883
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (8)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (28)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (30)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (47)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (873)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-892-877-73

Query Match 49.6%; Score 23.8; DB 10; Length 883;
Best Local Similarity 80.0%; Pred. No. 36;
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 7 TCCATCTCTGAGCCCTCTGCCATCTCTGTCACT 41
Db 701 TCTATCTCTCCAGCGGTATGTCTTCCCTGTCACT 735

RESULT 15
US-09-948-783-72
; Sequence 72, Application US/09948783
; Publication No. US20030100051A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et. al.
; TITLE OF INVENTION: 97 Human secreted proteins
; FILE REFERENCE: P2028P2
; CURRENT APPLICATION NUMBER: US/09/948,783
; CURRENT FILING DATE: 2001-09-10
```

; PRIOR APPLICATION NUMBER: 60/231,846  
 ; PRIOR FILING DATE: 2000-09-11  
 ; PRIOR APPLICATION NUMBER: 09/892,877  
 ; PRIOR FILING DATE: 2001-06-28  
 ; PRIOR APPLICATION NUMBER: 09/437,658  
 ; PRIOR FILING DATE: 1999-11-10  
 ; PRIOR APPLICATION NUMBER: PCT/US99/09847  
 ; PRIOR FILING DATE: 1999-05-06  
 ; PRIOR APPLICATION NUMBER: 60/085,093  
 ; PRIOR FILING DATE: 1998-05-12  
 ; PRIOR APPLICATION NUMBER: 60/085,094  
 ; PRIOR FILING DATE: 1998-05-12  
 ; PRIOR APPLICATION NUMBER: 60/085,105  
 ; PRIOR FILING DATE: 1998-05-12  
 ; PRIOR APPLICATION NUMBER: 60/085,180  
 ; PRIOR FILING DATE: 1998-05-12  
 ; PRIOR APPLICATION NUMBER: 60/085,927  
 ; PRIOR FILING DATE: 1998-05-18  
 ; PRIOR APPLICATION NUMBER: 60/085,906  
 ; PRIOR FILING DATE: 1998-05-18  
 ; PRIOR APPLICATION NUMBER: 60/085,924  
 ; PRIOR FILING DATE: 1998-05-18  
 ; PRIOR APPLICATION NUMBER: 60/085,922  
 ; PRIOR FILING DATE: 1998-05-18  
 ; PRIOR APPLICATION NUMBER: 60/085,921  
 ; PRIOR FILING DATE: 1998-05-18  
 ; PRIOR APPLICATION NUMBER: 60/085,923  
 ; PRIOR FILING DATE: 1998-05-18  
 ; PRIOR APPLICATION NUMBER: 60/085,925  
 ; PRIOR FILING DATE: 1998-05-18  
 ; PRIOR APPLICATION NUMBER: 60/085,928  
 ; PRIOR FILING DATE: 1998-05-18  
 ; PRIOR APPLICATION NUMBER: 60/085,920  
 ; PRIOR FILING DATE: 1998-05-18  
 ; NUMBER OF SEQ ID NOS: 465  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 72  
 ; LENGTH: 883  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: SITE  
 ; LOCATION: (8)  
 ; OTHER INFORMATION: n equals a,t,g, or c  
 ; NAME/KEY: SITE  
 ; LOCATION: (28)  
 ; OTHER INFORMATION: n equals a,t,g, or c  
 ; NAME/KEY: SITE  
 ; LOCATION: (30)  
 ; OTHER INFORMATION: n equals a,t,g, or c  
 ; NAME/KEY: SITE  
 ; LOCATION: (47)  
 ; OTHER INFORMATION: n equals a,t,g, or c  
 ; NAME/KEY: SITE  
 ; LOCATION: (873)  
 ; OTHER INFORMATION: n equals a,t,g, or c  
 ; US-09-948-783-72

Query Match 49.6%; Score 23.8; DB 10; Length 883;  
 Best Local Similarity 80.0%; Pred. No. 36;  
 Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 7 TCCATCTCTGCAGCCCTCTGCATCTCTGTCTACT 41  
 |||||  
 Db 701 TCTATCTCTCCAGCCGTATGTCTTCCCTGTCTACT 735

Search completed: April 28, 2005, 17:07:56  
 Job time : 93.6588 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 28, 2005, 03:05:59 ; Search time 195.782 Seconds  
(without alignments)  
9332.242 Million cell updates/sec

Title: US-10-075-846-3\_COPY\_574\_621

Perfect score: 48

Sequence: 1 agctcatcactactgcag.....catctctgcacttcaagtt 48

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext.1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: gb\_est1:\*

2: gb\_est2:\*

3: gb\_hic:\*

4: gb\_est3:\*

5: gb\_est4:\*

6: gb\_est5:\*

7: gb\_est6:\*

8: gb\_gss1:\*

9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
c 1	28.8	60.0	673	9	CE235158 tigr-gss-
c 2	27.8	57.9	802	9	CL694721 PRI0165a
c 3	26	54.2	231	4	BF948217 CM2-NN115
c 4	26	54.2	235	7	CV101020 UMC-pcrryb
c 5	26	54.2	264	7	CV101264 UMC-pcrryb
c 6	26	54.2	284	7	CV101330 UMC-pcrryb
c 7	26	54.2	265	7	CV101758 UMC-pcrryb
c 8	25.8	53.8	665	8	BZ724707 OGEDS19TM
c 9	25.8	53.8	853	8	BZ724703 OGEDS19TC
c 10	25.8	53.8	1337	9	CG744715 P037-2-F0
c 11	25.4	52.9	680	4	BI724543 1031072F0
c 12	25.4	52.9	804	9	BU651556 1112094C0
c 13	25.4	52.9	804	9	BU651556 1112094C0
c 14	25	52.1	266	1	AL835518 AL835518
c 15	25	52.1	358	1	AL842548 AL842548
c 16	25	52.1	422	1	AL836072 AL836072
c 17	25	52.1	426	1	AL835483 AL835483
c 18	25	52.1	430	1	AI332005 f966c12.y
c 19	25	52.1	441	1	AL835177 AL835177
c 20	25	52.1	526	1	AL834659 AL834659
c 21	25	52.1	547	8	AQ508245 RPCI-11-2
c 22	25	52.1	716	3	CR711148 Tetrarodon
c 23	25	52.1	806	7	CK467449 938753 MA
c 24	25	52.1	889	7	CK465348 936449 MA

25	24.8	51.7	490	6	CD739189
c 26	24.8	51.7	610	8	AZ386786
c 27	24.8	51.7	623	9	CR180890 Forward s
c 28	24.8	51.7	675	9	CR064241 Forward s
c 29	24.8	51.7	714	3	CR676321 Tetrarodon
c 30	24.8	51.7	860	9	CR002427 Forward s
c 31	24.8	51.7	894	9	BR957628 Forward s
c 32	24.8	51.7	970	9	CR179900 Forward s
c 33	24.6	51.3	276	2	BB159552 BB159552
c 34	24.6	51.3	402	9	CE052118 tigr-gss-
c 35	24.6	51.3	454	7	CF602831 BACCA01.0
c 36	24.6	51.3	586	7	CV093457 FAMU USDA
c 37	24.6	51.3	658	7	CNS47349 EST_15338
c 38	24.6	51.3	670	7	CF215885 CAST0002
c 39	24.6	51.3	710	6	CB685353 CSJNEF15K
c 40	24.6	51.3	748	9	CC580162 CH240_374
c 41	24.6	51.3	849	7	CF200536 RR89015N0
c 42	24.6	51.3	897	7	CF213803 CGF100080
c 43	24.4	50.8	329	7	CN236759 WUB113B03
c 44	24.4	50.8	483	1	AU260810 AU260810
c 45	24.4	50.8	584	7	CK943039 4066948 B

#### ALIGNMENTS

RESULT 1  
LOCUS CE235158/c 673 bp DNA linear GSS 25-SEP-2003  
DEFINITION tigr-gss-dog-1700033378690 Dog Library Canis familiaris genomic,  
genomic survey sequence.  
ACCESSION CE235158  
VERSION CE235158.1 GI:35390916  
KEYWORDS GSS.  
SOURCE Canis familiaris (dog)  
ORGANISM Canis familiaris  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
REFERENCE 1 (bases 1 to 673)  
AUTHORS Kirkness,E.F., Bafna,V., Halpern,A.L., Levy,S., Remington,K.,  
Ruskin,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and  
Venter,J.C.  
TITLE The dog genome: survey sequencing and comparative analysis  
JOURNAL Science 301 (5641), 1898-1903 (2003)  
MEDLINE 22875432  
PUBMED 14512627  
COMMENT Contact: Kirkness EF  
The Institute for Genomic Research  
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,  
Rockville, MD 20850, USA  
Tel: 301-838-0200  
Fax: 301-838-0208  
Email: ekirknes@tigr.org  
Class: shotgun.  
Location/Qualifiers  
source  
1. .673  
/organism="Canis familiaris"  
/mol\_type="genomic DNA"  
/strain="Standard Poodle"  
/db\_xref="taxon:9615"  
/clone\_lib="Dog Library"  
/note="Site 1: BstXI; Libraries were prepared from  
peripheral blood"

Query Match 60.0%; Score 28.8; DB 9; Length 673;  
Best Local Similarity 93.8%; Pred. No. 16;  
Matches 30; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 8 CCATCTCTGCAGCCCTCTGCCATCTCTGTCA 39  
DB 32 CCACATCTGCAGCCCTCTGCCATCTCTGTCA 1

JOURNAL  
MEDLINE  
PUBMED

/note="Vector: GTGTAACCTA; Funding: A grant from the Monsanto Company to the University of Missouri. Genetic Source: Day 6 pig blastocyst stage embryos were flushed from a gilt's reproductive tract in TL Hepes and cultured in NCSU23 (1) for 3 hours prior to vitrification. For vitrification, embryos were cultured in 1M dimethylsulfoxide and 1.3M ethylene glycol in Tissue Culture Medium (TCM) and 20% newborn calf serum for 5 minutes at 39 degrees Celsius. Then embryos were transferred to 3.2M ethylene glycol and 0.6M sucrose in TCM and 20% newborn calf serum for 45 to 60 seconds at 39 degrees Celsius. Capillary action was used to load the embryos in open pulled straws and the straws were plunged into liquid nitrogen. At thawing embryos were placed in 0.2M sucrose in TCM and 20% newborn calf serum for 1 minute, then 0.1M sucrose in TCM and 20% newborn calf serum for 5 minutes. Embryos were then rinsed and cultured in NCSU23 for 12 or 24 hours. Library Construction (PCR protocol): The amount of mRNA that was recovered from embryos was quite limiting and was not sufficient for library production with a standard protocol. Therefore, PCR-based protocol was utilized for producing libraries. The blastocysts were then rinsed with 50 ul polyvinylalcohol-PBS in depc treated medium three times and snap frozen with 1 ul super-RNasin in a 0.5ml RNase-free tube. Hybond messenger affinity paper was used to isolated the mRNA (2). The mRNA was reverse transcribed with a NotI-tag-dT18 oligonucleotide and a SMART oligonucleotide (Clontech) modified to contain a SalI site to generate full-length cDNA with a sequence complementary to the SMART oligonucleotide. Sequences within the SMART and dT oligonucleotides were used as primers to amplify the cDNAs by PCR with pfu turbo polymerase (Stratagene). The resulting PCR products were purified, digested with NotI and SalI and size fractionated by using a Chroma Spin-400 followed by a Chroma Spin-1000 column (Clontech). Purified cDNA from each PCR reaction was ligated into the pCMV-SPORT6 vector. Preliminary library characterization: Randomly chosen clones from each library were analyzed by restriction digestion to determine average insert size (96 clones) and by sequencing (~3 96-well plates) to confirm library quality (e.g. the presence of short polyA+ tails, genomic DNA contamination (must be <1%), ribosomal RNA clones (must be <1%), etc.) and to provide a sequence database representing the predominant clones in each library. The clones were sequenced at the University of Missouri-Columbia DNA Core Facility. Bioinformatics work was performed by GK Springer's bioinformatics group (WG Spollen, J Ries, A Gullin, AA Khambati, RV Patel, CM Topinka, SB Bhuiyan) in Computer Science and Health Management and Informatics Departments at the University of Missouri-Columbia. Clone Requests: Requests for clones should be made to the Director of the University of Missouri DNA Core facility at: porcine@net.missouri.edu. Citations: 1. Im G-S, Lai L, Liu Z, Hao Y and Prather RS. In vitro development of preimplantation porcine nuclear transfer embryos cultured in different media and gas atmospheres. Theriogenology 61: 1125-1135, 2004. 2. Anderson JE, Matteri RL, Abeydeera LR, Day BN and Prather RS. Degradation of maternal Cdc25c during the maternal to zygotic transition is dependent upon embryonic transcription. Molecular Reproduction & Development 60: 181-188, 2001.

TAG\_TISSUE=Blastocyst after cryopreservation  
TAG\_SEQ=GTGTAACCTA"

ORIGIN

Query Match 54.2%; Score 26; DB 7; Length 235;  
Best Local Similarity 85.3%; Pred. No. 1.3e+02;  
Matches 29; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

13 CTCTGAGCCCTCGACCTCTGTGCTTTCAG 46  
|||||

Db RESULT 5  
CV101264  
LOCUS  
DEFINITION  
CV101264 264 bp mRNA linear EST 09-SEP-2004  
scrofa cDNA 3', mRNA sequence.  
CV101264  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Sus scrofa (pig)  
Sus scrofa  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
1 (bases 1 to 264)  
WOODS,R., Whitworth,K., Springer,G.K., Forrester,L.J.,  
Spollen,W.G., Ries,J., Bivens,N., Murphy,C.N., Mathialigan,N.,  
Agca,Y., Green,J.A. and Prather,R.S.  
Swine Genomics  
Unpublished (2004)  
Contact: DNA Core Facility (Swine Project)  
Animal Science - RS Prather  
University of Missouri-Columbia  
M616 Medical Sciences Bldg., Columbia, MO 65212, USA  
Tel: (573)882-0428  
Fax: (573)884-5552  
Email: porcine@net.missouri.edu  
POLYA=Yes.

FEATURES  
source

Location/Qualifiers  
1..264  
/organism="Sus scrofa"  
/mol\_type="mRNA"  
/db\_xref="taxon:9823"  
/dev\_stage="Blastocyst after cryopreservation"  
/clone\_lib="pcrybst"  
/note="Vector: GTGTAACCTA; Funding: A grant from the Monsanto Company to the University of Missouri. Genetic Source: Day 6 pig blastocyst stage embryos were flushed from a gilt's reproductive tract in TL Hepes and cultured in NCSU23 (1) for 3 hours prior to vitrification. For vitrification, embryos were cultured in 1M dimethylsulfoxide and 1.3M ethylene glycol in Tissue Culture Medium (TCM) and 20% newborn calf serum for 5 minutes at 39 degrees Celsius. Then embryos were transferred to 3.2M ethylene glycol and 0.6M sucrose in TCM and 20% newborn calf serum for 45 to 60 seconds at 39 degrees Celsius. Capillary action was used to load the embryos in open pulled straws and the straws were plunged into liquid nitrogen. At thawing embryos were placed in 0.2M sucrose in TCM and 20% newborn calf serum for 1 minute, then 0.1M sucrose in TCM and 20% newborn calf serum for 5 minutes. Embryos were then rinsed and cultured in NCSU23 for 12 or 24 hours. Library Construction (PCR protocol): The amount of mRNA that was recovered from embryos was quite limiting and was not sufficient for library production with a standard protocol. Therefore, PCR-based protocol was utilized for producing libraries. The blastocysts were then rinsed with 50 ul polyvinylalcohol-PBS in depc treated medium three times and snap frozen with 1 ul super-RNasin in a 0.5ml RNase-free tube. Hybond messenger affinity paper was used to isolated the mRNA (2). The mRNA was reverse transcribed with a NotI-tag-dT18 oligonucleotide and a SMART oligonucleotide (Clontech) modified to contain a SalI site to generate full-length cDNA with a sequence complementary to the SMART oligonucleotide. Sequences within the SMART and dT oligonucleotides were used as primers to amplify the cDNAs by PCR with pfu turbo polymerase (Stratagene). The resulting PCR products were purified, digested with NotI and SalI and size fractionated by using a Chroma Spin-400 followed by a Chroma Spin-1000 column (Clontech). Purified cDNA from each PCR reaction was ligated into the

pcmv-sport6 vector. Preliminary Library Characterization: Randomly chosen clones from each library were analyzed by restriction digestion to determine average insert size (96 clones) and by sequencing (~3 96-well plates) to confirm library quality [e.g. the presence of short polyA+ tails, genomic DNA contamination (must be <1%), ribosomal RNA clones (must be <1%), etc.] and to provide a sequence database representing the predominant clones in each library. The clones were sequenced at the University of Missouri-Columbia DNA Core Facility. Bioinformatics work was performed by GK Springer's bioinformatics group (WG Spollen, J Ries, A Gullin, AA Khambati, RV Patel, CM Topinka, SB Bhuiyan) in Computer Science and Health Management and Informatics Departments at the University of Missouri-Columbia. Clone Requests: Requests for clones should be made to the Director of the University of Missouri DNA Core facility at: porcine@net.missouri.edu. Citations: 1. Im G-S, Lai L, Liu Z, Hao Y and Prather RS. In vitro development of preimplantation porcine nuclear transfer embryos cultured in different media and gas atmospheres. Theriogenology 61: 1125-1135, 2004. 2. Anderson JE, Matteri RL, Abeydeera LR, Day BN and Prather RS. Degradation of maternal Cdc25c during the maternal to zygotic transition is dependent upon embryonic transcription. Molecular Reproduction & Development 60: 181-188, 2001.

TAG TISSUE=Blasocyst after cryopreservation

TAG\_SEQ=GTGTAACTTA

#### ORIGIN

Query Match 54.2%; Score 26; DB 7; Length 264;  
Best Local Similarity 85.3%; Pred. No. 1.4e+02;  
Matches 29; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 13 CTCTGAGCCCTCTGCCACTCTGTCTACCTTCAG 46  
DB 111 CTCTGAGCCCTCTGCCACTCTGTCTACCTTCAG 144

#### RESULT 6

LOCUS CV101330 264 bp mRNA linear EST 09-SEP-2004  
DEFINITION UMC-pcrybst1-008-d10 Blastocyst after cryopreservation pcryst Sus  
scrofa cDNA 3', mRNA sequence.

ACCESSION CV101330

VERSION CV101330.1 GI:51583494

KEYWORDS EST.

SOURCE Sus scrofa (pig)

#### ORGANISM

Sus scrofa  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

REFERENCE 1 (bases 1 to 264)

AUTHORS Wood, R., Whitworth, K., Springer, G.K., Forrester, L.J.,  
Spollen, W.G., Ries, J., Bivens, N., Murphy, C.N., Mathialigan, N.,  
Agca, Y., Green, J.A. and Prather, R.S.

TITLE Swine Genomics

JOURNAL Unpublished (2004)

COMMENT Contact: DNA Core Facility (Swine Project)

Animal Science - RS Prather

University of Missouri-Columbia

M616 Medical Sciences Bldg., Columbia, MO 65212, USA

Tel: (573)882-0428

Fax: (573)884-5552

Email: porcine@net.missouri.edu

POLYA=Yes.

#### FEATURES

source

1..264

/organism="Sus scrofa"

/mol\_type="mRNA"

/db\_xref="taxon:9823"

/dev\_stage="Blastocyst after cryopreservation"

/clone\_lib="pcrybst"

/note="Vector: GTGTAACTTA; Funding: A grant from the

Monsanto Company to the University of Missouri. Genetic Source: Day 6 pig blastocyst stage embryos were flushed from a gilt's reproductive tract in TL Hapes and cultured in NGSU3 (1) for 3 hours prior to vitrification. For vitrification, embryos were cultured in 1M dimethylsulfoxide and 1.3M ethylene glycol in Tissue Culture Medium (TCM) and 20% newborn calf serum for 5 minutes at 39 degrees Celsius. Then embryos were transferred to 3.2M ethylene glycol and 0.6M sucrose in TCM and 20% newborn calf serum for 45 to 60 seconds at 39 degrees Celsius. Capillary action was used to load the embryos in open pulled straws and the straws were placed in into liquid nitrogen. At thawing embryos were placed in 0.2M sucrose in TCM and 20% newborn calf serum for 1 minute, then 0.1M sucrose in TCM and 20% newborn calf serum for 5 minutes. Embryos were then rinsed and cultured in NGSU3 for 12 or 24 hours. Library Construction (PCR Protocol): The amount of mRNA that was recovered from embryos was quite limiting and was not sufficient for library production with a standard protocol. Therefore, PCR-based protocol was utilized for producing libraries. The blastocysts were then rinsed with 50 ul polyvinylalcohol-PBS in depc treated medium three times and snap frozen with 1 ul super-RNasin in a 0.5ml

Knase-free tube. Hybrid messenger affinity paper was used to isolate the mRNA (2). The mRNA was reverse transcribed with a NotI-tag-dT18 oligonucleotide and a SMART oligonucleotide (Clontech) modified to contain a SalI site to generate full-length cDNA with a sequence complementary to the SMART oligonucleotide. Sequences within the SMART and dr oligonucleotides were used as primers to amplify the cDNAs by PCR with pfu turbo polymerase (Stratagene). The resulting PCR products were purified, digested with NotI and SalI and size fractionated by using a Chroma Spin-400 followed by a Chroma Spin-1000 column (Clontech). Purified cDNA from each PCR reaction was ligated into the pcMV-SPORT6 vector. Preliminary Library Characterization: Randomly chosen clones from each library were analyzed by restriction digestion to determine average insert size (96 clones) and by sequencing (~3 96-well plates) to confirm library quality [e.g. the presence of short polyA+ tails, genomic DNA contamination (must be <1%), ribosomal RNA clones (must be <1%), etc.] and to provide a sequence database representing the predominant clones in each library. The clones were sequenced at the University of Missouri-Columbia DNA Core Facility. Bioinformatics work was performed by GK Springer's bioinformatics group (WG Spollen, J Ries, A Gullin, AA Khambati, RV Patel, CM Topinka, SB Bhuiyan) in Computer Science and Health Management and Informatics Departments at the University of Missouri-Columbia. Clone Requests: Requests for clones should be made to the Director of the University of Missouri DNA Core facility at: porcine@net.missouri.edu. Citations: 1. Im G-S, Lai L, Liu Z, Hao Y and Prather RS. In vitro development of preimplantation porcine nuclear transfer embryos cultured in different media and gas atmospheres. Theriogenology 61: 1125-1135, 2004. 2. Anderson JE, Matteri RL, Abeydeera LR, Day BN and Prather RS. Degradation of maternal Cdc25c during the maternal to zygotic transition is dependent upon embryonic transcription. Molecular Reproduction & Development 60: 181-188, 2001.

TAG TISSUE=Blasocyst after cryopreservation

TAG\_SEQ=GTGTAACTTA

#### ORIGIN

Query Match 54.2%; Score 26; DB 7; Length 264;  
Best Local Similarity 85.3%; Pred. No. 1.4e+02;  
Matches 29; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 13 CTCTGAGCCCTCTGCCACTCTGTCTACCTTCAG 46  
DB 111 CTCTGAGCCCTCTGCCACTCTGTCTACCTTCAG 144





```

/clone_lib="ZM_0.7_1.5_KB"
/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match      53.8%; Score 25.8; DB 8; Length 665;
Best Local Similarity 73.3%; Pred. No. 1.9e+02;
Matches 33; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 3 CTCATCCACTCTGCAGCCCTCTGCCATCTCTGTCACTTCAGT 47
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 577 CTCATCCCTACTTTCGGCGCGTCGCACATCTCTATCAAAATTCAT 533

RESULT 9
BZ724703
LOCUS      853 bp      DNA      linear      GSS 03-MAR-2003
DEFINITION      OGEDS19TC ZM_0.7_1.5_KB Zea mays genomic clone ZMBMa0253C13,
genomic survey sequence.
ACCESSION      BZ724703
VERSION        BZ724703.1 GI:28697951
KEYWORDS       GSS.
SOURCE         Zea mays
ORGANISM       Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 853)
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
Other_GSSs: OGEDS19TM
Contact: Cathy Whitelaw
TIgR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
Location/Qualifiers
1. .853
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZMBMa0253C13"
/clone_lib="ZM_0.7_1.5_KB"
/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match      53.8%; Score 25.8; DB 8; Length 853;
Best Local Similarity 73.3%; Pred. No. 2e+02;
Matches 33; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 3 CTCATCCACTCTGCAGCCCTCTGCCATCTCTGTCACTTCAGT 47
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 376 CTCATCCCTACTTTCGGCGCGTCGCACATCTCTATCAAAATTCAT 420

RESULT 10
CG744715/c
LOCUS      1337 bp      DNA      linear      GSS 24-OCT-2003
DEFINITION      P037-2-F02.ya Ppa EcoRI BAC Library Pristionchus pacificus genomic,
genomic survey sequence.
ACCESSION      CG744715
VERSION        CG744715.1 GI:37965583
KEYWORDS       GSS.
SOURCE         Pristionchus pacificus
ORGANISM       Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Neodiplogasteridae; Pristionchus.
1 (bases 1 to 1337)
Srinivasan, J., Sinz, W., Jesse, T., Wiggers-Perebolte, L., Jansen, K.,
Buntjer, J., van der Meulen, M. and Sommer, R.J.
An integrated physical and genetic map of the nematode Pristionchus
pacificus
Mol. Genet. Genomics 269 (5), 715-722 (2003)
12884007
Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: raif.sommer@tuebingen.mpg.de
Class: BAC ends.
Location/Qualifiers
1. .1337
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Ppa EcoRI BAC Library"
/notes="The library was generated by a partial digest of
the genomic DNA with EcoRI and cloning into the BAC
vector."

ORIGIN
Query Match      53.8%; Score 25.8; DB 9; Length 1337;
Best Local Similarity 73.3%; Pred. No. 2.2e+02;
Matches 33; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 4 TCATCCACTCTGCAGCCCTCTGCCATCTCTGTCACTTCAGT 48
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 851 TCATTCATCATCATCATCTCTCTCTCTCTCTCTCTCTCAATT 807

RESULT 11
BZ724543/c
LOCUS      680 bp      mRNA      linear      EST 19-SEP-2001
DEFINITION      1031072F07.y1 C. reinhardtii CC-1690, Stress II (normalized),
Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION      BZ724543
VERSION        BZ724543.1 GI:15700238
KEYWORDS       EST.
SOURCE         Chlamydomonas reinhardtii
ORGANISM       Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaeae; Chlamydomonas.
1 (bases 1 to 680)
Grossman, A., Chang, C.-W., Davies, J., Harris, E., Hauser, C.,
Lefebvre, P., McDermott, J.P., Shrager, J., Silflow, C. and Sern, D.
Analyses of the Chlamydomonas reinhardtii Genome: A Model,
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants. Project: 1031
Unpublished (2001)
Contact: Charles Hauser
DCMB Box 91000
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu
Location/Qualifiers
1. .680
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Stress II
(normalized), Lambda Zap II"

```

/note="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2: XhoI; Stress condition II library, constructed by John Davies and Jeffrey McDermott, combines cDNAs from CC-1690 cells grown to mid-log phase in TAP (NH<sub>4</sub><sup>+</sup> - containing) and shifted to TAP - NO<sub>3</sub><sup>-</sup> (24hrs); H2 production conditions (0, 12hr, 24hr) see Melis et al., (2000) Plant Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP + sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr). PolyA mRNA was purified from each sample, pooled and synthesized. The cDNA was directionally cloned into lambda Zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites. pBluescript II SK- plasmids were excised from the lambda Zap clones by superinfection with ExAssist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al., (1996) Genome Research 6: 791-806."

## ORIGIN

Query Match 52.9%; Score 25.4; DB 4; Length 680;  
Best Local Similarity 74.4%; Pred. No. 2.7e+02;  
Matches 32; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 GCTCATCCATCTCTGCAGCCCTCTGCCATCTCTGTCTCTTC 44  
|||||  
Db 337 GCGGTCCACCATTCGAGCCCGTGCATCTCTGTCTCTTC 295

## RESULT 12

BU651556/c

LOCUS 716 bp mRNA linear EST 30-SEP-2002  
DEFINITION 1112094C02.y1 C. reinhardtii CC-1690 (mt+), CC-1691 (mt-), Gamete (normalized), Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.

ACCESSION BU651556

VERSION 1

KEYWORDS EST.

SOURCE Chlamydomonas reinhardtii

ORGANISM Chlamydomonas reinhardtii

Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;

Chlamydomonadaceae; Chlamydomonas.

REFERENCE 1 (bases 1 to 716)

AUTHORS Groseman, A., Chang, C.-W., Davies, J., Harris, E., Hauser, C.,

Lefebvre, P., McDermott, J.P., Shrager, J., Silflow, C. and Stern, D.

Analyses of the Chlamydomonas reinhardtii Genome: A Model,

Unicellular System for Analyzing Gene Function and Regulation in

Vascular Plants. Project: 1112

Unpublished (2002)

CONTACT: Charles Hauser

DCMB Box 91000

Duke University

Durham, NC 27708-1000

Tel: 919 613 8159

Fax: 919 613 8177

Email: chauser@duke.edu.

Location/Qualifiers

1. 716

/organism="Chlamydomonas reinhardtii"

/mol\_type="mRNA"

/strain="21gr (CC-1690 wild type mt+) &amp; 6145C (CC-1691

wild type mt-)"

/db\_xref="taxon:3055"

/clone\_lib="C. reinhardtii CC-1690 (mt+), CC-1691 (mt-),

Gamete (normalized), Lambda Zap II"

/note="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2:

XhoI; Gamete library was constructed by Hui Zhao, Min Lu,

Jeffrey McDermott, William J. Snell and John Davies.

Strain 21gr cells (CC-1690; mating type plus) and strain

6145C cells (CC-1691; mating type minus) that had been

growing on a light-dark cycle (13:11 L/D) in R-medium

(Sager and Granick) were separately transferred into

nitrogen-free medium at 8 hours into the light period.

PolyA mRNA was purified from each sample every 2 hours for

the next 18 hours. The mRNA was pooled and used for cDNA

synthesis. The cDNA was directionally cloned into lambda Zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites. pBluescript II SK- plasmids were excised from the lambda Zap clones by superinfection with ExAssist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al., (1996) Genome Research 6: 791-806."

## ORIGIN

Query Match 52.9%; Score 25.4; DB 5; Length 716;  
Best Local Similarity 74.4%; Pred. No. 2.7e+02;  
Matches 32; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 GCTCATCCATCTCTGCAGCCCTCTGCCATCTCTGTCTCTTC 44  
|||||  
Db 198 GCGGTCCACCATTCGAGCCCGTGCATCTCTGTCTCTTC 156

## RESULT 13

CW521717

LOCUS 804 bp DNA linear GSS 08-OCT-2004

DEFINITION OP\_Ba0022J02.f OP\_Ba Oryza punctata genomic clone OP\_Ba0022J02

5', genomic survey sequence.

ACCESSION CW521717

VERSION 1

KEYWORDS GSS.

SOURCE Oryza punctata

ORGANISM Oryza punctata

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoidae; Oryzaceae; Oryza.

REFERENCE 1 (bases 1 to 804)

AUTHORS SanMiguel, P., Westerman, R., Kim, H., Yu, Y., Wissotski, M., Yost, D.,

Stum, D., Rao, K., Luo, M., Jetty, R., Kudrna, D., Muller, C.,

Hatfield, J., Soderlund, C., Wing, R. and Jackson, S.A.

OMAP Project - Purdue University

Unpublished (2004)

CONTACT: Scott A. Jackson

Jackson Laboratory

Purdue University

915 W. State St., West Lafayette, IN 47907, USA

Tel: 7654963621

Fax: 7654967255

Email: sjackson@purdue.edu

Basecalling by phred version 0.020425.c. This sequence was derived

from the raw sequence read by clipping with lacy version 1.19a.

Bases 42-845 of the raw sequence (length 1357) were retained after

clipping.

PCR Primers

FORWARD: TAA TAC GAC TCA CTA TAG GG

BACKWARD: CAC TCA TTA GGC ACC CCA

Insert Length: 161000 Std Error: 0.00

Plate: 0022 row: J column: 02

Seq primer: TAA TAC GAC TCA CTA TAG GG

Class: BAC ends.

Location/Qualifiers

1. 804

/organism="Oryza punctata"

/mol\_type="genomic DNA"

/db\_xref="taxon:4537"

/clone="OP\_Ba0022J02"

/tissue\_type="young leaves"

/lab\_host="DH10B-T1 phage resistant"

/clone\_lib="OP\_Ba"

/note="Vector: PAGIBAC1; Site\_1: HindIII; Site\_2: HindIII"

```

Db      214 GCTCATCCATCTCCGCGCCCTTGTGTCATCTGCATTTTC 256

RESULT 14
LOCUS   AL835518/c
DEFINITION
Takifugu rubripes (Fugu rubripes)
EST.
AL835518.1 GI:21877480
ACCESSION
AL835518
VERSION
AL835518
KEYWORDS
Takifugu rubripes (Fugu rubripes)
SOURCE
Takifugu rubripes
ORGANISM
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Takifugu.
1 (bases 1 to 266)
Clark, M.S., Edwards, Y.J., Peterson, D., Clifton, S.W., Thompson, A.J.,
Sasaki, M., Suzuki, Y., Kikuchi, K., Watabe, S., Kawakami, K.,
Sugano, S., Elgar, G. and Johnson, S.L.
Fugu ESTs: new resources for transcription analysis and genome
annotation
Genome Res. 13 (12), 2747-2753 (2003)
Contact: Clark MS
MRC Human Genome Mapping Project Resource Centre
Hinxtun, Cambridge, CB10 1SB, UK
Email: biohelp@hgmrc.ac.uk
Vector: pBlue-script II KS
V type: phagemid
PRIMER: KS
The clone can be obtained from www.hgmrc.ac.uk
Library created by Greg Elgar
MRC Human Genome Mapping Project Resource Centre, Hinxtun,
Cambridge, CB10 1SB, UK
Library sequenced by Sarah Warner and Jim Hills
MRC Human Genome Mapping Project Resource Centre, Hinxtun,
Cambridge, CB10 1SB, UK
FEATURES             source
    source
    1..266
        /location=Qualifiers
        /organism="Takifugu rubripes"
        /mol_type="mRNA"
        /db_xref="taxon:31033"
        /clone="EPR025apck3"
        /tissue_type="spleen"
        /clone_lib="EPR0"
        /note="Vector: pBlue-script II KS"
ORIGIN
Query Match      52.1%; Score 25; DB 1; Length 266;
Best Local Similarity 75.6%; Pred. No. 3.1e+02;
Matches 31; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY      3 CTCATCCATCTCGAGCCCTCTGCCATCTCTGTCACTTT 43
        |||||
Db      167 CTCATCCAGATCAGCAGCTCTCTGGCAGCTTTGCGAATT 127

Query Match      52.1%; Score 25; DB 1; Length 266;
Best Local Similarity 75.6%; Pred. No. 3.1e+02;
Matches 31; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY      3 CTCATCCATCTCGAGCCCTCTGCCATCTCTGTCACTTT 43
        |||||
Db      167 CTCATCCAGATCAGCAGCTCTCTGGCAGCTTTGCGAATT 127

Query Match      52.1%; Score 25; DB 1; Length 266;
Best Local Similarity 75.6%; Pred. No. 3.1e+02;
Matches 31; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY      3 CTCATCCATCTCGAGCCCTCTGCCATCTCTGTCACTTT 43
        |||||
Db      167 CTCATCCAGATCAGCAGCTCTCTGGCAGCTTTGCGAATT 127

Query Match      52.1%; Score 25; DB 1; Length 358;
Best Local Similarity 75.6%; Pred. No. 3.3e+02;
Matches 31; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY      3 CTCATCCATCTCGAGCCCTCTGCCATCTCTGTCACTTT 43
        |||||
Db      199 CTCATCCAGATCAGCAGCTCTCTGGCAGCTTTGCGAATT 159

Search completed: April 28, 2005, 16:13:12
Job time : 202.782 secs

REFERENCE
AUTHORS
Clark, M.S., Edwards, Y.J., Peterson, D., Clifton, S.W., Thompson, A.J.,
Sasaki, M., Suzuki, Y., Kikuchi, K., Watabe, S., Kawakami, K.,
Sugano, S., Elgar, G. and Johnson, S.L.
Fugu ESTs: new resources for transcription analysis and genome
annotation
Genome Res. 13 (12), 2747-2753 (2003)
Contact: Clark MS
MRC Human Genome Mapping Project Resource Centre
Hinxtun, Cambridge, CB10 1SB, UK
Email: biohelp@hgmrc.ac.uk
Vector: pME18S-FL3
V type: phagemid
PRIMER: ME-735FW
Library created by Koichi Kawakami, Masahide Sasaki, Yutaka Suzuki,
Sumio Sugano
The Institute of Medical Science, The University of Tokyo,
Shirokanedai, Minato-ku, Tokyo 108-8639, Japan
and
Kiyoshi Kikuchi, Shugo Watabe
Laboratory of Aquatic Molecular Biology and Biotechnology, Graduate
School of Agricultural and Life Sciences, The University of Tokyo,
Bunkyo-ku, Tokyo 108-8639, Japan
Library sequenced by Melody S. Clark and Amanda Thompson MRC Human
Genome Mapping Project Resource Centre Hinxtun, Cambridge, CB10
1SB, UK.
FEATURES             Location/Qualifiers
    source
    1..358
        /organism="Takifugu rubripes"
        /mol_type="mRNA"
        /db_xref="taxon:31033"
        /clone="F000H06aF1"
        /sex="female"
        /tissue_type="heart"
        /dev_stage="adult"
        /clone_lib="F000H"
        /note="Vector: pME18S-FL3"
ORIGIN
Query Match      52.1%; Score 25; DB 1; Length 358;
Best Local Similarity 75.6%; Pred. No. 3.3e+02;
Matches 31; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY      3 CTCATCCATCTCGAGCCCTCTGCCATCTCTGTCACTTT 43
        |||||
Db      199 CTCATCCAGATCAGCAGCTCTCTGGCAGCTTTGCGAATT 159

Search completed: April 28, 2005, 16:13:12
Job time : 202.782 secs

```